

THIS WEEK

EDITORIALS

JAPAN US and European physicists should send next collider east **p.312**

WORLD VIEW The danger of a successful strategy for funds **p.313**

SPRAY DAY Tanzanian toads sent home to freshly sprinkled gorge **p.316**



The limits of free speech

Unregulated drug marketing stifles science and harms patients. To suggest otherwise is an affront to liberty — not a protection of it.

On 12 December, American drug giant Pfizer agreed to pay the US government more than US\$55 million in fines. Wyeth, a pharmaceutical company now owned by Pfizer, had overstated the benefits of its proton-pump inhibitor Protonix (pantoprazole), despite repeated warnings from the US Food and Drug Administration (FDA).

A government prosecutor said that the exaggerations were intentional and planned: “Wyeth tried to cheat the system by obtaining a limited FDA approval for Protonix, fully intending to promote this drug for additional, unapproved uses.”

Collectively, drug companies have paid billions of dollars for similar offences. Again and again, the lure of increased sales overcomes the threat of potential penalties.

Another court ruling this month could tip the balance further, and not in the FDA's favour. This case, *United States v. Caronia*, relates to the conviction of a pharmaceutical salesman for Orphan Medical (taken over in 2005 by Jazz Pharmaceuticals based in Dublin), who was recorded telling a potential client that the drug Xyrem (γ-hydroxybutyrate) could treat a variety of ills, including muscle disorders, chronic pain, Parkinson's disease and fibromyalgia. In fact, the drug had been approved by the FDA only for treating acute muscle weakness or excessive daytime sleepiness associated with narcolepsy. Xyrem can be abused and carries the FDA's strongest safety warning against use in children. In 2007, Jazz Pharmaceuticals agreed to a settlement of \$20 million in fines and compensation payments for the off-label marketing of Xyrem by Orphan.

However, a three-judge federal appeals court has now overturned the salesman's misdemeanour conviction, arguing two to one that off-label promotion should be protected as free speech. If the case continues all the way to the Supreme Court, and it probably will, it could have a profound effect on how drugs are marketed in the United States. Even under the existing regulations, drug companies face regular accusations that they promote illnesses that don't exist, craft clinical trials so as to show their drugs in the best light rather than to answer questions of efficacy and safety, and bury evidence about dangerous side effects.

Some in the field predict that to protect off-label marketing as free speech would unleash a barrage of television ads and expose patients to unnecessary, even fatal, side effects. “It's a potential catastrophe for patients, and I don't use that term lightly,” says Steven Nissen, head of the Cleveland Clinic in Ohio.

At first glance, the free-speech argument has merit. If a pharmaceutical company wants to market a drug, it must first convince the FDA that it is safe and effective for a specific use. Once a drug has FDA approval, however, doctors are generally free to prescribe it as they choose. Indeed, ‘off-label’ prescriptions for medical uses without FDA approval are common: some doctors do prescribe Xyrem for complaints other than narcolepsy. If drugs work as doctors expect, both patients and pharmaceutical companies win. Why, then, should the United States continue to prohibit sales representatives from

informing physicians of other possible uses for their drugs?

Off-label uses are off-label precisely because the drugs have not been studied well enough to understand their benefits and risks in this context. Doctors can learn about off-label uses and their drawbacks from talking to colleagues, attending conferences and reading the medical literature. These sources of information are influenced to some extent by drug companies' marketing policies, but they are clearly less biased than direct promotion to doctors by company salespeople.

Moreover, prohibitions on off-label promotion encourage good science. At present, companies must invest in further clinical trials to convince the FDA to broaden a drug's use. If the FDA cannot regulate how drugs are marketed, the incentive to conduct such trials will evaporate.

Sometimes the investment pays off. Johnson & Johnson, based in New Brunswick, New Jersey, recently announced an FDA-approved label expansion for its anti-testosterone drug Zytiga (abiraterone) to patients with advanced prostate cancer who have not already received chemotherapy. This could significantly increase Zytiga's sales.

But trials can also show that broadening the label is not justified. London-based AstraZeneca conducted a two-year, 1,300-patient trial hoping to show that its anticholesterol drug Crestor was superior to a competitor's. But Crestor did not perform better than Pfizer's Lipitor in preventing plaque build-up in arteries. In another case, large clinical trials revealed that two anti-arrhythmia drugs thought to protect against heart attacks actually increased death rates.

A weaker FDA with less authority to restrict off-label marketing could leave drug companies with even fewer incentives to show that off-label therapies work. This would harm patients and discourage good science. ■

A burden weighed

Despite some shortcomings, a global study of health metrics should be applauded.

The best evidence-based health policies are made on the basis of thorough and regular updates of the global burden of disease. Aid agencies need to know how and where to target their funding and monitor impact. Knowledge of what sickens and kills people can help health-care providers and researchers to tailor their priorities to needs and to long-term trends in the health of populations. Yet in many poor areas of the world, basic systems such as death certification are lacking, and quality health-care data are scarce and scattered.

At national, regional and global levels, collection of comprehensive health metrics is too often the poor cousin in health care and in health research. The result is a shortage of high-quality, comparable and readily accessible data.

For these reasons, the publication in *The Lancet* last week of the Global Burden of Disease 2010 study (GBD 2010) should be applauded. The study is an unprecedented five-year effort by hundreds of researchers in dozens of countries worldwide, who made a huge effort to track down, collate and analyse surveys, as well as published and unpublished health-care data. Led by the Institute for Health Metrics and Evaluation at the University of Washington in Seattle, they have produced a vast smorgasbord of global estimates of the burden of multiple diseases, injuries, risk factors and chronic complications (see page 322).

The findings capture the world's health status in impressive detail, and highlight significant trends, including how, in almost all countries, life expectancies are rapidly converging towards the long lives previously enjoyed only by the richest. The study also charts the demise of major causes of global health burden such as infectious diseases and maternal and child mortality, although these continue to blight many poorer countries, particularly in sub-Saharan Africa.

People benefit from longer lives but they are increasingly experiencing a downside. They spend many of those extra years in ill health, often with more than one illness, which creates large costs for health-care systems. The study found that age-related illnesses include not only the usual suspects such as cancers and heart disease, but also a host of conditions that rarely kill but often disable, such as mental illness and musculoskeletal disorders.

Health-care systems and research agendas must adapt accordingly; for example, health-care providers must learn how to manage the

high costs of tackling this new disease landscape. The GBD 2010 should be required reading for health leaders and research administrators, and should lead them to re-examine how well current research portfolios match emerging trends in the burden of sickness and disease.

The GBD 2010 is far from perfect. Some of the underlying data are weak — for example, they may be scarce, unreliable or unstandardized

“The GBD 2010 should be required reading for health leaders and research administrators.”

— and in places the study relies heavily on sophisticated modelling to eke out meaning. To its credit, it has tried to provide transparency in the form of quantitative indicators as to the robustness — or otherwise — of each of its estimates. Given such caveats, some of its findings will inevitably be dubious. Malaria researchers have vigorously contested the GBD's assertion earlier this year

that malaria killed twice as many people in 2010 as previously thought, but the study stands by its data (see *Nature* <http://doi.org/j2s>; 2012).

Such disagreements are inevitable in any complex, large-scale, international collaborative undertaking of this kind, particularly in areas where uncertainties in the data are highest. Such health metrics are also highly political, because they can affect the direction of national and international health-care and research funds, where there is much turf to fight over. But, ultimately, the findings of the GBD 2010, however imperfect, provide a robust basis and analytical framework for further research and health policies that is better than anything that went before. That its findings and methodologies will be challenged and debated in the months and years to come is not only healthy, but how science works. The way forward is to rework and build on the foundations that have been laid. ■

Head of the line

Japanese scientists deserve support in their bid for the next big collider.

The International Linear Collider (ILC) is, literally, a pipe dream in the minds of physicists. The proposed 31-kilometre-long superconducting machine is the heir apparent to the Large Hadron Collider (LHC) — the world's most powerful particle accelerator — at CERN, Europe's particle-physics laboratory near Geneva, Switzerland. For now, the ILC remains a paper study, but particle physicists would like to see it built so that they can probe ever more fundamental questions about the nature of matter.

During the best of times, large, multibillion-dollar facilities such as the ILC take decades of planning and political preparation, and these are not the best of times. As the world's governments struggle through the worst financial crisis in a generation, the prospects for any project of this scale are daunting to say the least.

But there is one possibility: Japan. The island nation has yearned for a big international project such as the ILC, and it very nearly got one in 2005, but narrowly lost out to France on the giant ITER fusion project. Politicians and physicists there who had backed the ITER bid were dejected but determined, and they have formed a strong, cross-party political group to bring the ILC to Japan.

This coalition has worked diligently for years to promote a Japanese bid, but two developments make it even more likely to happen. First, on 11 March 2011, a devastating earthquake and tsunami rocked the Tohoku region of Japan. Reconstruction yen have been flowing since, and a science city being built in the region has emerged as a possible engine for development (a second site for the collider is also on the table, in the southern region of Kyushu). Second, the discovery of

the Higgs boson at CERN in July provides the ILC with a scientific target — even running at less than full energy, the ILC could provide a detailed study of the new particle.

Taking account of these two events, politicians in both of Japan's major political parties seem increasingly interested in an ILC-type project. It even formed part of the party platform of the Liberal Democratic Party, which won last weekend's parliamentary elections. Moreover, Japan is technically ready for the project. A slew of high-profile experiments built in the 1990s and 2000s show that it has the skill and industrial know-how to take on an advanced accelerator. It has also been leading ILC research and development in several areas, including crucial work on how to focus the beams.

But Japan can't do it alone. It will need expertise, funds and in-kind contributions from around the globe. Europe seems to be coming around to the idea. After all, Japan has contributed to CERN, and the lab is busy with its own collider. In the United States, support is considerably softer. The main US laboratory, Fermilab in Batavia, Illinois, recently lost its main accelerator and has begun an ambitious neutrino programme in its stead. Budgets are not sufficient to support that, the ongoing LHC work and the ILC.

US and European scientists should throw their support behind the ILC project. For the forward-thinking Europeans, this means a clear commitment to the parts and manpower they might be able to supply. For the Americans, it would probably require a willingness to slow the neutrino programme. That may be hard for them to swallow but, deep down, US physicists know that participation in the ILC is the only real option if the nation is to remain at the vanguard of particle physics.

Statements of support from overseas will not guarantee that the ILC will go ahead. Japan's opaque government will still have to debate the programme's merits internally and come up with a process for committing to its construction. But an early show of support could give the collider the push it needs to get under way. That would be a great victory for Japan, and the world. ■

➤ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunqy



Time to stop relying on things past

US science advocates are depending on strategies and statistics that may not survive contemporary politics, says David Goldston.

In times of fiscal crisis, the reflex response of interest groups is to circle the wagons and defend the status quo, arguing that any cuts would be fatal. Predictably, the scientific community in the United States is currently following that script, and it is a perfectly reasonable and well executed Plan A. But there is not much of a Plan B, and there should be, especially given that the United States could be entering a prolonged period of relative austerity.

It is not surprising that science advocates feel comfortable with, and even complacent about, Plan A. They have been making essentially the same arguments for basic research for nearly 70 years — linking it to economic growth and other public benefits — and few in politics have challenged them. Moreover, the administration of US President Barack Obama is egging them on, as part of its effort to ratchet up pressure on the Republican-controlled House of Representatives to raise revenues instead of slashing spending. It's hard to have doubts about Plan A when a natural target — those who propose the budget — urges you to implement it more fiercely.

Yet the time-honoured can also be the shopworn, and scientists and their advocates should address the weaknesses in this strategy before the pitfalls become more visible.

The first problem is that some of Plan A's familiar claims cannot survive close scrutiny. For example, two prestigious groups, the President's Council of Advisors on Science and Technology and the National Academy of Sciences, this year issued reports that call for total US research-and-development spending to increase until it reaches at least 3% of gross domestic product. The Obama administration has cited this goal for some time, and the European Union has a similar goal.

Unfortunately, there is little analytical basis for this number; it is a political assertion masquerading as a mathematical constant. The only argument in its favour is that some nations spend that much. Those countries hardly add up to a coherent case; they include nations with smaller economies, as well as countries such as Japan, which has been in the economic doldrums for years.

The number does not take into account what kind of research is being pursued, or by what sectors of the economy, or how much is actually being spent. A dissertation could be written on how 3% acquired its magical properties, but the number doesn't even seem to be based on spending levels in some past golden age. Rather, it is more of a back-of-the-envelope comparison done when the United States got nervous about competitiveness in the 1980s.

Is there any real harm in this? Yes, in at least two ways. The target helps to fuel the scientific community's demoralizing and distorted sense that it is being neglected and wilfully financed at suboptimal levels. And it leaves science advocates unable to answer persuasively if officials ask for proof that

funding is inadequate or for an estimate of what would be sufficient.

Science lobbyists have begun to address a second weakness in Plan A: its insularity. Scientists often talk as though science funding is decided in a vacuum, yet a major determinant of the research budget is the overall level of civilian domestic spending. The vacuum-packed approach generally made sense when overall spending was relatively stable. Why get involved in high-stakes, partisan feuding over the budget when it is safer to stand on the sidelines?

For more than a year, though, the House Republicans and Obama have engaged in a fundamental battle over the shape of the federal budget, and all domestic spending is imperilled without new tax revenues. So science groups have taken tentative steps towards weighing in on broader issues. In a letter to Congress and Obama earlier this month, more than 100 science and university groups said that tax revenues and reform of programmes such as Medicare should be part of any deficit agreement, without endorsing any specific proposal. This strategy can alienate some politicians and so carries risks. And in some ways, it's a shame to see every group sucked into the political maelstrom. But the change in strategy makes sense; to sit out a debate that will set the context for your future is even riskier.

That kind of strategic calculus should lead advocates to confront a graver weakness in their reliance on Plan A. No matter what agreement the President and the Republicans reach, it is likely to usher in, at best, an era of slower growth for science spending. Thanks to previous deals, federal

spending is already expected to be flat in real terms for the next decade. If scientists fail to plan for that eventuality, then decisions will be made without them. As part of plan B, they need to think about questions such as: which agencies, programmes and fields are the lowest priorities? How can the government ensure that younger researchers are still able to get funding? Is there a thoughtful way to shrink a system that produces ever more grant proposals?

The scientific community must also devise better ways to figure out when spending cuts are actually causing serious problems. Simple measures such as increased proposal pressure will not convince. It also needs ways to evaluate specific cost-cutting strategies.

Plan A has seen science advocates through some close scrapes, and Obama officials say that they will try to protect science funding. But the plan pretty much assumes that fiscal policy and politics are not fundamentally different today from what they were in the 1950s, and that success just demands more intensive lobbying. That's too risky a notion not to have a back-up. ■

David Goldston is director of government affairs at the Natural Resources Defense Council in Washington DC. Views are his own. e-mail: partyofonecolumn@gmail.com

**THERE IS NOT
MUCH OF A
PLAN B,
AND THERE
SHOULD BE.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/54tafy

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

ASTRONOMY

From the Kuiper Belt to comets

A hunt for the smallest members of the Kuiper Belt — a disk of icy cold objects at the Solar System's edge — has revealed a potential source of local comets.

Hilke Schlichting at the University of California, Los Angeles, and her colleagues made use of the Hubble Space Telescope's Fine Guidance Sensors to search the Kuiper Belt. These sensors stabilize the telescope by watching distant stars, which are occasionally eclipsed by a passing Kuiper Belt Object. By trawling more than nine years' worth of data, the team found a single candidate for a new object just 530 metres across. Combining that result with a previous one, the team estimates that small objects in the Kuiper Belt are abundant enough to be the source of the short-period comets observed in the inner Solar System.

Astrophys. J. 761, 150 (2012)

PALAEOECOLOGY

Teeth speak of dietary change

The ecosystems of the Omo Valley in southern Ethiopia changed about 2.8 million years ago, causing a shift in the diet of herbivores towards C4 grasses. Such grasses exhibit a more efficient photosynthetic pathway than their more common C3 cousins.

Faysal Bibi of the Natural History Museum in Berlin and his colleagues found increased amounts of the isotope carbon-13 in the tooth enamel of two herbivores, the antelope *Tragelaphus nakuae* and the pig-like *Kolpochoerus limnetes*, after the 2.8-million-year mark. This points to a diet high in C4



ANIMAL BEHAVIOUR

Benefits of mixed flocks

Cooperation between species may be an overlooked factor in community organization.

Ecologists tend to assume that all of the species within a community are more different from one another than chance would predict, because competition precludes two species from sharing the same ecological niche. So mixed-species bird flocks are a puzzle: these groups, which contain birds of different species that all eat roughly the same food, seem to show species happily co-existing. According to an analysis of a global data set by Kartik Shanker at the Indian Institute of Science in Bangalore and his co-authors, the more similar two birds are by taxonomy, body size and foraging style, the more likely they are to be found together, especially where tropical mixed flocks are concerned.

The team suggests that this happens because cooperative benefits — such as those provided by alarm calls or the discovery of food resources — are best obtained from the most similar individuals.

Am. Nat. 180, 777–790 (2012)

plants, which accumulate more carbon-13 than do C3 plants.

The findings tally with evidence of a major environmental shift in Africa at the time, the authors say. However, they add that it is not yet clear whether the incoming C4 grasses replaced woodlands or C3 grasses.

Biol. Lett. <http://dx.doi.org/10.1098/rsbl.2012.0890> (2012)

ENVIRONMENTAL SCIENCE

Rivers' antibiotic resistance threat

Synthetic antibiotic-resistant genes have found their way into microorganisms in Chinese rivers.

Min Jin and Jun-Wen Li at the Key Laboratory of Risk Assessment and Control for Environment and Food Safety in Tianjin, China, and their colleagues extracted microbes from the water of six rivers. They then looked for lab-made DNA fragments called plasmid vectors, which are used as genetic tools to introduce and express particular genes in cells. In all six rivers, the authors found microbes containing genes from plasmid vectors widely used in biotechnology, as well as synthetic versions of the β -lactam antibiotic (*bla*) gene, which confers resistance to ampicillin.

The team warns that this plasmid-containing lab waste could be a source of animal and human antibiotic resistance.

Environ. Sci. Technol. <http://dx.doi.org/10.1021/es302760s> (2012)

ECOLOGY

When plants run the food chain

Some plants may offer snacks to attract carnivorous insects that provide a defence against herbivorous attackers.

RANGU NARAYAN

SAMUEL BECK



Billy Krimmel at the University of California, Davis, and Ian Pearse at Cornell University in Ithaca, New York, report that sticky, hair-like trichomes on common tarweed (*Mardia elegans*) trap small insects. These may function as bait to attract predatory insects (pictured) that feast on caterpillars, which would damage the plant.

The authors carried out a field experiment in California, and found that adding dead fruitflies to plants increased the abundance of the predators that they surveyed by between 76% and 450%. Adding bait also reduced bud damage from the most common caterpillar in the ecosystem by 60%, and increased fruit production by 10%.

Ecol. Lett. <http://dx.doi.org/10.1111/ele.12032> (2012)

GENETICS

RNA tails time protein production

A cell's rhythm of protein production may be regulated by the addition and removal of tails at the end of the messenger RNA molecules that carry the protein recipes.

The cellular circadian clock yields rhythmic patterns in the expression of many genes and

the synthesis of their protein products. Carla Green and her colleagues at the University of Texas Southwestern Medical Center in Dallas found 237 mouse-liver mRNAs that show circadian rhythmicity in the number of adenosine molecules added to the end of mRNAs. The length of this 'poly(A) tail' can affect the stability of the mRNA and, ultimately, the amount of cognate protein produced.

Further analysis of five of these mRNAs showed that peak poly(A) tail length preceded a peak in protein accumulation by 4–8 hours, even when the amount of those mRNAs remained constant.

Genes Dev. 26, 2724–2736 (2012)

MATERIALS

Grains call a halt with sound

Granular materials such as sand may slow high-speed projectiles by turning kinetic energy from an incoming object into complex networks of sound waves.

Abram Clark at Duke University in Durham, North Carolina, and his colleagues studied the impact of bronze disks falling onto a bed of photoelastic grains — plastic particles that light up under pressure. Filming at 40,000 frames per second, the team found that falling disks pulsed acoustic energy along branching networks of grains that glowed as pressure travelled through the material (pictured). The 'force chains' decayed exponentially and died out after passing through about 10 particles. Falling disks decelerated in spurts that

COMMUNITY CHOICE

The most viewed papers in science

TAXONOMY

Scarce cetaceans catalogued

HIGHLY READ
on www.cell.com
16 Nov–16 Dec

A paper titled "The world's rarest whale" documents the unprecedented beaching of two spade-toothed beaked whales (*Mesoplodon traversii*).

Members of this species, which haunt the deepest reaches of the South Pacific Ocean, had been known from only a tiny handful of bones and teeth. A mother and calf stranded in 2010 are thought to be the first to be seen in the flesh. The adult was about five metres long and patterned in black, white and grey. Initially misidentified as Gray's beaked whales (*M. grayi*), the rare animals were correctly identified thanks to DNA typing by Rochelle Constantine at the University of Auckland in New Zealand and her co-authors.

Curr. Biol. 22, R905–R906 (2012)

matched closely in time with acoustic fluctuations through the grain bed, suggesting that acoustic waves play a key part in dissipating the energy of projectiles that strike granular materials.

Phys. Rev. Lett. 109, 238302 (2012)

GEOLOGY

Counting geoneutrinos

Neutrinos are tiny, nearly massless particles that are generated by nuclear reactions and can fly unimpeded through even the thickest rock. Because of these properties, geologists would like to use geoneutrinos — electron antineutrinos emitted during the decay of natural radionuclides — to estimate the heat generated by radioactivity inside Earth's mantle.

Ondřej Šrámek at the University of Maryland in College Park and his colleagues now present a plan for making measurements of these particles at two or more strategic points in the Pacific Ocean to discriminate between different models of Earth's mantle architecture.

Earth Planet. Sci. Lett. <http://dx.doi.org/10.1016/j.epsl.2012.11.001> (2012)

CANCER

Clonal clues reveal cancer chaos

The general assumption that a tumour cell's behaviour is guided largely by its DNA sequence may be incorrect.

John Dick at the University of Toronto in Canada and his colleagues grew 150 cell populations from individual, labelled cancer cells taken from ten patient-derived colorectal-cancer samples. The researchers transplanted these cells into mice, then tracked their genetics and behaviour as the grafts grew. Tumour-cell clones from the same cell line differed widely in their growth patterns, ability to seed new tumours and response to the cancer drug oxaliplatin.

The results suggest that factors beyond the DNA sequence — such as differences in the chemical groups attached to it, which influence gene expression, or environmental influences — shape tumour-cell conduct.

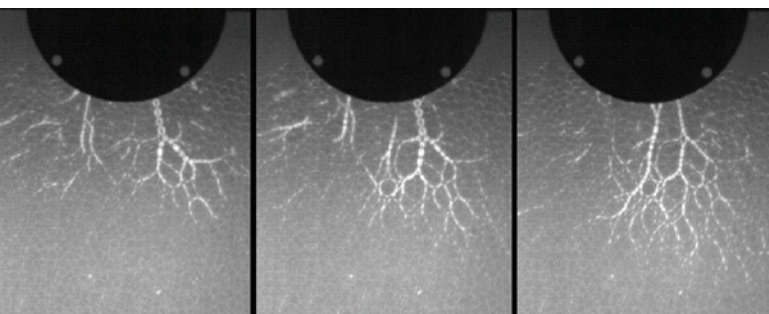
Science <http://dx.doi.org/10.1126/science.1227670> (2012)

NATURE.COM

For the latest research published by Nature visit:

www.nature.com/latestresearch

ABRAM CLARKE AND ROBERT BEHRINGER



SEVEN DAYS

The news in brief

RESEARCH

Moon smash

Twin spacecraft that mapped the gravity field of the Moon with unprecedented precision ended their mission with a controlled crash on 17 December. Ebb and Flow, the two probes that make up NASA's Gravity Recovery and Interior Laboratory (GRAIL), produced the first ultra-high-resolution picture of the Moon's gravitational field by flying in tandem in lunar orbit for just under a year while exchanging radio signals. See go.nature.com/uoa1ft for more.

Ice plunge halted

A broken boiler has scuppered the first attempt to drill through more than 3 kilometres of ice to reach Lake Ellsworth, an ancient subglacial lake in Antarctica. The main burner-control circuit blew as a team from the British Antarctic Survey powered up its hot-water drill on 10 December. The team fitted a secondary burner to keep the boiler going until a replacement part could arrive, but that, too, packed up four to five days later. The team is unlikely to resume drilling before 21 December.

Collisions pause

CERN, Europe's particle-physics laboratory near Geneva, Switzerland, announced on 17 December that it had stopped its run of proton-proton collisions at the Large Hadron Collider, in which the Higgs boson was discovered. For about a month in early 2013, the accelerator will instead smash protons with lead ions to recreate conditions just after the Big Bang. It will then shut down until 2015 for an overhaul that will increase the proton-collision energies by more than 60%.



JULIE LARSEN MAHER/WILDLIFE CONSERVATION SOCIETY

Sprinklers aid a toad's return

A toad declared extinct in the wild in 2009 has been reintroduced into its original location in Tanzania, the Wildlife Conservation Society's Bronx Zoo announced on 12 December. The construction in 1999 of a hydroelectric dam had damaged the original waterfall microhabitat

of the Kihansi spray toad (*Nectophrynoides asperginis*) in the Kihansi Gorge. While naturalists bred the toad in captivity, the Tanzanian government built an artificial spray environment in the gorge. That enabled the zoo and its partners to return 2,000 toads to the wild.

Ancient ice

Australia will lead a team to Antarctica next summer to collect an ice core spanning the past 2,000 years, environment minister Tony Burke announced on 15 December. The 400-metre core, drilled from an area of high snowfall called Aurora Basin North, should yield a year-by-year picture of climate changes over that period.

eLife launches

The open-access journal *eLife* officially launched on 13 December. The online publication, for advances in life and biomedical sciences, has been publishing articles since October, but it now has a dedicated website. The journal, which currently charges no author fees, is the brainchild of three major funding organizations: the

US Howard Hughes Medical Institute, Germany's Max Planck Society and the UK Wellcome Trust.

PEOPLE

Negligence found

Germany's main funding agency, the DFG, has prohibited immunologist Silvia Bulfone-Paus from submitting funding proposals for three years. An investigating committee had said Bulfone-Paus had committed a 'gross negligence of her supervisory duty' after it found that two members of her lab at the Research Center Borstel had manipulated data in four papers. See go.nature.com/wed2en for more.

Oceans chief quits

Marine ecologist Jane Lubchenco announced her resignation as head of the

US National Oceanic and Atmospheric Administration (NOAA) on 12 December. During her tenure at NOAA, Lubchenco pushed through a national ocean policy as well as scientific-integrity rules for the agency. She had previously said that she would be willing to stay on for President Barack Obama's second term, but in a memo to staff, Lubchenco cited family issues in choosing to return to Oregon, where she is a professor at Oregon State University in Corvallis, at the end of February.

POLICY

Texas cancer probe

Texas prosecutors have launched a criminal probe into irregularities in grant review at a US\$3-billion publicly funded cancer agency. On 7 December, the district attorney in Travis

County, home to the Cancer Prevention and Research Institute of Texas (CPRIT) in Austin, notified executive director William Gimson of the investigation. Gimson resigned on 10 December. On the same day, CPRIT announced the appointment of a new chief scientific officer: cancer immunologist Margaret Kripke.

Animal ruling

A higher administrative court in Bremen, Germany, ruled on 11 December that local health authorities had been wrong in 2008 when they chose not to reapprove a licence for studies on macaque monkeys granted to neuroscientist Andreas Kreiter of the University of Bremen. The authorities' decision had followed animal-rights protests. But the court ruled that licensing decisions should be made by expert committees, in line with European laws. See go.nature.com/w2dwzd for more.

Ivory seized

About 1,500 pieces of African elephant ivory have been seized in Malaysia, its customs authority revealed last week. Unconfirmed reports suggest that the stash weighs up to 24 tonnes, which would make it the largest haul of illegal ivory ever. The ivory (pictured) was hidden in ten crates being shipped from Togo to China,



and the seizure has reignited concerns over the growing illegal trade. See go.nature.com/oik71p for more.

EU patent package

After decades of negotiation, the European Union (EU) approved a unified 'patent package' on 11 December. The agreement will allow applicants to obtain a unitary patent that will be valid in 25 out of the 27 EU member states, after Spain and Italy refused to sign up to the deal. The first unitary patents are expected to be granted in April 2014. See go.nature.com/jgyvly for more.

Drugs neglect

Of the 756 drugs approved from 2000 to 2011, 3.8% were for 'neglected diseases' that afflict poor countries, according to a report from the aid organization Médecins Sans Frontières and the Drugs for Neglected Diseases Initiative, a Swiss not-for-profit research and development programme. Just 1.1% of 1,393 approvals in the previous

25 years were for such diseases. The findings were presented last week. See go.nature.com/nvwhoi for more.

Safety for sharks

The Cook Islands has joined forces with neighbouring French Polynesia to create the world's largest shark sanctuary, an area of 6.7 million square kilometres. Roughly the size of Australia, the sanctuary has a ban on shark fishing and possession or sale of shark products. French Polynesia established most of the area on 6 December, and the Cook Islands added another 1.9 million square kilometres on 13 December.

BUSINESS

India's drug law

India adopted legislation to cap the price of essential medicines last week — and promptly faced a challenge in its Supreme Court. The All India Drug Action Network filed a petition against the measure, an update of a 17-year-old law. Although the goal of the legislation is to make medicines more affordable, observers say that loopholes in the policy could allow pharmaceutical companies to sidestep the price ceiling. Critics say that poor people might end up paying higher prices for some drugs. See go.nature.com/irsqur for more.

COMING UP

1 JANUARY

California launches its market for trading carbon emissions — the second largest in the world, behind only the European system. See go.nature.com/8xfsqx

2 JANUARY

Barring a last-minute deal, legislation that makes US federal science agencies subject to automatic 'sequestration cuts' takes effect, removing around 8% from their programme budgets. See go.nature.com/7nrwj

IP superpower

China's State Intellectual Property Office (SIPO) handled 526,412 patent applications in 2011, surpassing the 503,582 received by the US Patent and Trademark Office, according to a report published on 11 December by the World Intellectual Property Organization in Geneva, Switzerland. This makes SIPO for the first time the world's largest patent office. See page 323 for more.

EVENTS

Troubled rocket

North Korea may finally have achieved its space dream — only to see its satellite spiral out of control. The country has been successful in its fifth attempt since 1998 to launch an object into orbit, as confirmed by the US–Canadian defence radar network NORAD on 11 December. However, as *Nature* went to press, US reports suggest that the spinning satellite is 'dead' in orbit — with no transmissions.

➔ NATURE.COM

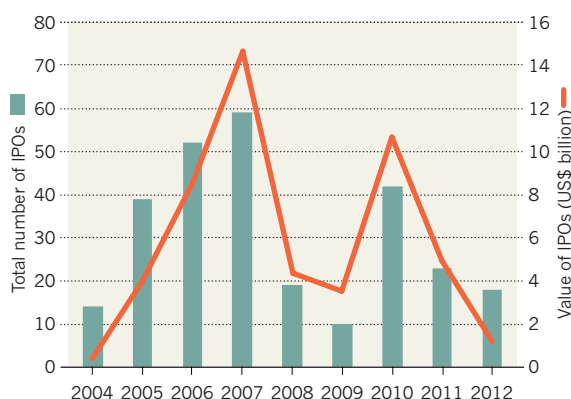
For daily news updates see: www.nature.com/news

TREND WATCH

It has been a dismal year for initial public offerings (IPOs) of clean-technology firms (see chart). "Clean energy has been hit with a double whammy of oversupply and declining policy support," says Stefan Linder, an analyst at Bloomberg New Energy Finance in New York. The largest debut was by Jiangsu Sunrain Solar Energy in Lianyungang, China, which raised US\$340 million. Most recently, on 13 December, solar installer SolarCity of San Mateo, California, raised \$92 million in its IPO.

CLEAN-TECH IPO CRUNCH

Initial public offerings (IPOs) in clean energy suffered a dismal year in 2012, raising the lowest total since 2004.



NEWS IN FOCUS

HIGHER EDUCATION From Russia, tough love for universities **p.320**

BUSINESS Pioneering genetics firm gets a new lease of life **p.321**

INTELLECTUAL PROPERTY China's patent boom sparks legal battles **p.323**

2012 IN REVIEW The news, the people, the images **p.324**



R. DURÁN/ESO/NAOJ/NRAO/ALMA



The Atacama Large Millimeter/submillimeter Array (ALMA), set on a 5,000-metre-high plateau in Chile, will ultimately consist of 66 dish antennas.

ASTRONOMY

Mega-array reveals birthplace of giant stars

Early results from Atacama telescope signal the opening of a scientific frontier.

BY ERIC HAND IN PUERTO VARAS, CHILE

Astronomers have punched through a brick wall. They have seen into the heart of a massive cloud of cold dust and gas near the Galactic Centre that, because it blocks visible light, had been dubbed the 'brick'. Viewed in millimetre waves (which have wavelengths between those of microwaves and radio waves), the faintly glowing dust reveals knots of gas — the embryos of stars, seen in such detail that they could show how the Galaxy's most massive stars are born.

The images testify to the penetrating vision of astronomy's newest global megaproject, the Atacama Large Millimeter/submillimeter Array (ALMA) in Chile. "Up till now we've just been looking at big blobs," says Jill Rathborne of the Commonwealth Scientific and Industrial Research Organization's Astronomy and Space Science division in Sydney, Australia. She likens the new images to seeing the details

of a tree. "Now we're looking at the limbs, the branches, the flowers and the roots."

Rathborne presented her team's images, produced with just six hours of observing time, on 12 December here at the first ALMA science conference. The US\$1.4-billion array, perched on a 5,000-metre-high plateau near the Chilean-Bolivian border, is due to have all 66 of its dish antennas completed by the end of 2013. At that point, ALMA will be able to separate its movable antennas by as much as 16 kilometres, which will allow its signals to be integrated into images of the cold Universe that will have exquisite spatial and spectral resolution. But even at partial strength — Rathborne's work was done with 25 antennas — ALMA has reached a scientific frontier. That is a relief to director Thijs de Graauw. "We have been working 30 years to get here," he says.

➔ **NATURE.COM**
To learn more
about ALMA,
see:
go.nature.com/ogulno

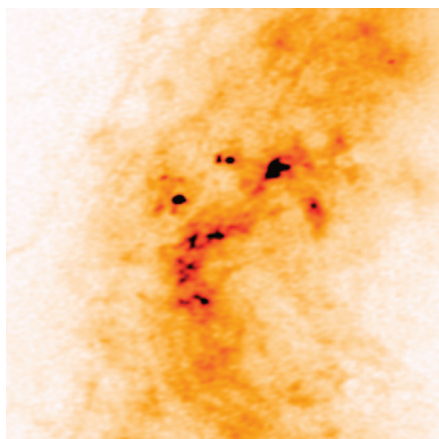
The array — an international collaboration involving North America, Europe, Taiwan and Japan — is unveiling cool, dusty parts of the Universe that are invisible at other wavelengths. ALMA astronomers looking outside the Milky Way are detecting bubbles of gas being expelled from distant galaxies — presumably propelled by jets shooting out from supermassive black holes at the galaxies' centres. The loss of gas in these bubbles is thought to put a brake on galaxy growth.

In our own Galaxy, ALMA is training its antennas on disks of gas and dust around young stars — material that eventually coalesces into planets. Jes Jørgensen of the Niels Bohr Institute at the University of Copenhagen presented the first discovery of glycolaldehyde around a protostar. This organic molecule is an essential building block of ribose, a component of RNA. "We have to get organic molecules from somewhere," says Neal Evans, an ALMA board member from the University of Texas at ►

► Austin. The ALMA find is a hint that they formed at the same time as the planets.

But with the array not yet at full strength, many of the main conference results came from large, nearby objects — in particular, the giant clumps of gas within our Galaxy that become clusters of stars. Astronomers especially want to identify the elusive conditions that lead to the growth of high-mass stars — those 8–150 times heavier than the Sun. Such behemoths are rare, but are thought to have been the Universe's first stars and to have had a major effect on cosmic history. Their intense ultraviolet light would have ionized the hydrogen of interstellar space, for example, and they would have quickly exploded as supernovae, seeding the Universe with elements heavier than hydrogen and helium.

Yet astronomers have never witnessed a high-mass star being born, and hotly debate how they form. The turbulent-core theory posits that they originate in an exceptionally large knot of dense gas — a 'core' — in which turbulent winds



ALMA sees star-forming gas clumps in the 'brick'.

boost the internal pressure. That would allow the core to grow bigger than it ordinarily would before igniting by thermonuclear fusion and producing an outgoing 'stellar wind' of material that blows away the remaining gas and cuts

off growth. A rival theory, known as competitive accretion, suggests that the cores break up within the gas cloud, forming small protostars that later become giants by eating up material elsewhere in the cloud before their stellar winds can blow the gas away.

Advocates of both theories are looking to natural laboratories such as the brick for answers. Rathborne has identified some 50 possible cores within the cloud, which contains material with an overall mass of 100,000 Suns. But the images are not quite sharp enough to show whether the cores are large and bounded — as the turbulent-core theory suggests — or are seeds in the process of breaking up into smaller cores, which competitive accretion calls for. That detail could start coming in January with the next cycle of ALMA experiments, which will be conducted with at least 32 antennas. Whatever they show, Rathborne says, the region is a lode that astronomers will be mining for quite some time. Bricks like this, she says, "are made of gold". ■

JILL RATHBORNE/CSIRO/JCMT/ALMA

HIGHER EDUCATION

Russia shakes up its universities

Government plans to close struggling institutions and increase funding to the best.

BY QUIRIN SCHIERMEIER

The New Year is unlikely to be a happy one for thousands of Russian university teachers and students whose institutes are facing massive cuts and closures in 2013, following a controversial performance review. But some researchers are optimistic that despite the hardships, the most significant overhaul of Russia's university system in living memory will help to improve science and innovation in the country.

The move reflects the determination of Russian authorities to end support for hundreds of under-achieving institutions and to concentrate funding in a smaller number of high-performing universities. Since the collapse of the Soviet Union, demand for academic degrees has soared, and the number of public and private universities has doubled, to around 1,100. But Russia's science output has not increased accordingly, and higher-education experts and employers have long voiced concerns over the poor quality of many university programmes. Insiders suggest that no more than 50 Russian institutions are up to international standards.

After his election as president in March, Vladimir Putin decreed an overhaul of higher education while promising to increase university funding gradually over the next decade (see *Nature* 483, 253–254; 2012). To identify weak universities, the Ministry of Science and

Education commissioned an external audit of almost 600 public higher-education institutions. The results, leaked last month, made for depressing reading. Almost 500 of the institutions — 102 universities and 374 local branches — were found wanting, on the basis of criteria such as the quality of students, research intensity and productivity, and the amount of teaching space. About 40 of Russia's top universities, already classed as elite institutions by the government, were not included in the review.

Twenty institutions, including the Moscow State University for the Humanities (MSUH) and the Moscow State Evening Metallurgical Institute, were found to be so severely below par that the auditors recommended that they should be either closed or merged with more proficient institutions. Around 100 other universities are to be maintained but need to "optimize" their teaching and research performance, the auditors said. The ministry has already asked these institutions to submit development plans outlining how they intend to improve their performance, and a decision will be made about their future in April.

The audit has created a stir among Russian academics. Critics say that niche universities such as the MSUH — which the reviewers have labelled 'ineffective' — concentrate on teaching rather than

scientific research, so it was unfair to judge them on research performance. "One should have been more thoughtful in designing the criteria," says Isak Froumin, a senior education specialist with the Higher School of Economics (HSE) in Moscow, and former leader of the World Bank education programme in Russia.

But several analysts and researchers contacted by *Nature* agree that science and innovation are likely to benefit as the reforms free up money to strengthen programmes in the surviving universities. "Russia needs better-trained university graduates and it needs more and stronger university research," says Leonid Gokhberg, first vice-rector of the HSE and head of its Institute for Statistical Studies and Economics of Knowledge. "I think the proposed overhaul will be helpful in both respects."

An educational bill likely to come into effect early next year should cement the reforms: it aims to reduce the number of universities even further. Meanwhile, Putin has promised a marked increase in academic salaries, as well as bonuses for special achievements in teaching and research. The Russian government also plans to carry out an audit in 2013 of the academic performance of private universities.

"Despite its shortcomings, this was an overdue exercise," says Froumin. "The outcome may be painful for some universities but what counts more is that Russian students have a right to receive a decent education." ■

► **NATURE.COM**
Read *Nature's* Global
Science special at:
nature.com/global

GENOMICS

Big biotech buys iconic genetics firm

Amgen's deal with deCODE Genetics shows value of combined medical and genetic data.

BY MONYA BAKER

More than a decade after the sizzle of genomics investments turned to a fizzle, a biotechnology giant is buying one of the field's flagship firms. The announcement on 10 December that Amgen, based in Thousand Oaks, California, will pay US\$415 million for Iceland's deCODE Genetics, a company with a rocky history, shows that drug developers are again ready to invest in genetic data to find therapeutic targets.

When genomics largely failed to deliver on the promise of new drug targets during the 2000s, many investors, including large drug firms, shed their stakes in companies that dealt with disease-related data and shifted to those with actual compounds in clinical trials. But data and analysis are precisely what deCODE Genetics, based in Reykjavik, has to offer. It has no drug candidates in active clinical trials, and is known chiefly for producing a steady flow of publications that pinpoint genetic culprits — and possible disease mechanisms — for conditions including neurodegeneration, cancer, cardiovascular disease and psychosis. For example, deCODE scientists this summer identified a genetic variant that protects against Alzheimer's disease (T. Jonsson *et al.* *Nature* **488**, 96–99; 2012).

The key to deCODE's scientific success is its access to a trove of genetic data, genealogies and medical records from some 140,000 Icelanders — roughly half of Iceland's population. Being able to correlate genetic information with Iceland's extensive medical records is one reason why deCODE has made discoveries that other firms have not, says John Bell, a geneticist at the University of Oxford, UK, who chairs a UK advisory group on human genomics. "The genetics bit is the easiest bit; it's the clinical data that has historically been the problem," he says.

The rise of low-cost sequencing and electronic medical records is allowing more information to be extracted from population-genetics studies, making genomics more attractive to drug companies. Indeed, deCODE's data, technology and expertise will help Amgen to identify which experimental drugs will be most likely to succeed, says Sean Harper, head of research and development at Amgen. Already, two of Amgen's drug candidates — for heart disease and osteoporosis — take aim at protein



TO THE BRINK AND BACK

Rescued from bankruptcy, Iceland's deCODE Genetics has found new life.

1996 deCODE Genetics founded by Kári Stefánsson (pictured).

2000 deCODE goes public, raising US\$173 million.

2004–05 deCODE announces clinical trials of drugs to prevent heart attacks and asthma.

2009 deCODE declares bankruptcy.

2010 deCODE bought by some of its original investors for \$14 million.

2012 Amgen announces it will acquire deCODE.

targets that were discovered using genetic data, says Harper. "We've become increasingly interested in having more examples of targets discovered or validated by human genetics."

By becoming part of a much bigger biotech firm, deCODE and its chief executive, Kári Stefánsson, may finally be able to fulfil a promise made throughout the company's 16-year history: to use genetic data to benefit patients. Stefánsson says that deCODE can "contribute to more health benefits faster with a company with a proven track record to make therapeutics".

Stefánsson, a gruff man who traces his roots

to the Vikings, left a professorship at Harvard Medical School in Boston, Massachusetts, and returned to his homeland to found deCODE in 1996. Like other companies from the genomics boom of the early 2000s, deCODE turned to drug development as a potential source of revenue, launching its own clinical trials in the mid-2000s (see 'To the brink and back'). But the small firm soon racked up huge debts and failed to win investors as Iceland was plunged into the 2008 financial crisis. The company declared bankruptcy in 2009 and was rescued by some of its original investors, who acquired the firm and its assets for around \$14 million and went on to invest another roughly \$45 million. "Even though we are [venture capitalists], it's not just about financial returns," says Terry McGuire, a co-founder of Polaris Venture Partners in Waltham, Massachusetts. "We decided it really was a world treasure."

After the deCODE acquisition, due to be finalized in the next two weeks, Stefánsson will stay on as president of deCODE and become a vice-president at Amgen. He says that deCODE will have no lay-offs and will probably even hire staff. Stefánsson emphasizes that the new ownership will not affect how Icelanders' data are managed. The DNA samples will remain in Iceland and access will be through deCODE, subject to existing privacy-protection policies and oversight by an ethics committee of Icelandic citizens.

Scientists from deCODE are some of the most-cited geneticists, and some researchers worry that as the Icelandic group boosts Amgen's product pipelines, deCODE's leadership in basic research will fade. "deCODE may find it far harder to continue publishing its findings if they're being treated as a potential competitive advantage," says Daniel MacArthur, a geneticist at Massachusetts General Hospital in Boston.

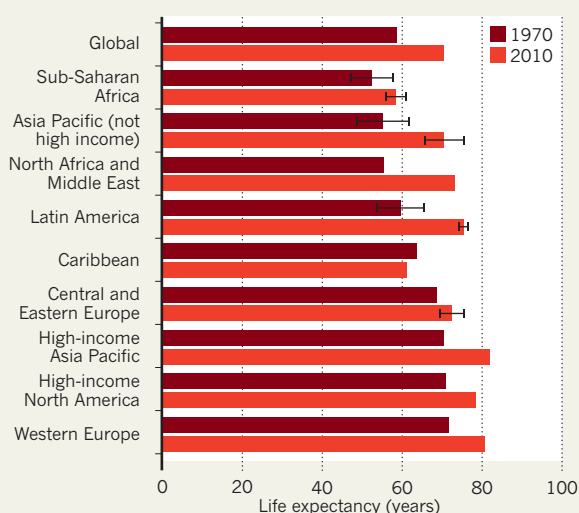
But Harper dismisses such concerns, saying that it is in Amgen's interest to keep deCODE's scientific edge razor sharp. "For us to in some way squelch that creativity would be very short-sighted." Stefánsson is confident that business at deCODE will continue as usual. "I am unafraid, unconcerned and enthusiastic about what is happening," he says. "We will continue to indulge in our discovery orgy." ■

NATURE.COM
The Encyclopedia of
DNA Elements:
nature.com/encode

REMY STEINER/REUTERS

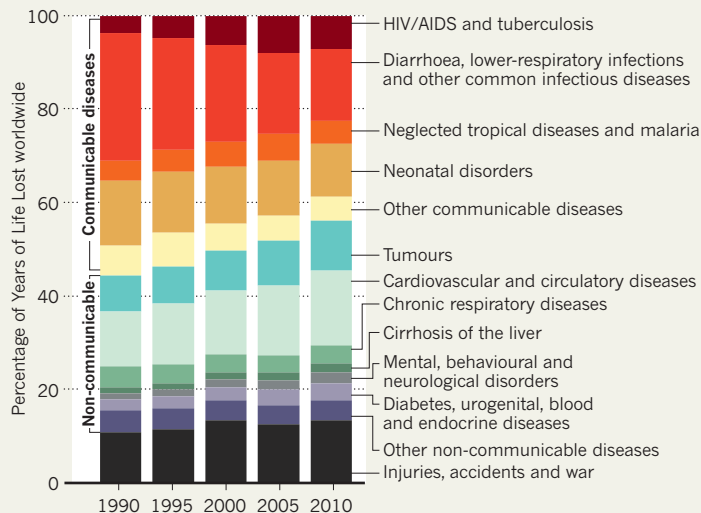
A SHIFTING PICTURE OF GLOBAL HEALTH

Life expectancies have shown extraordinary increases in many nations over the past 40 years. This, along with significant declines in communicable, maternal and neonatal diseases, means that many countries must adapt to focus more on the health needs of ageing populations.

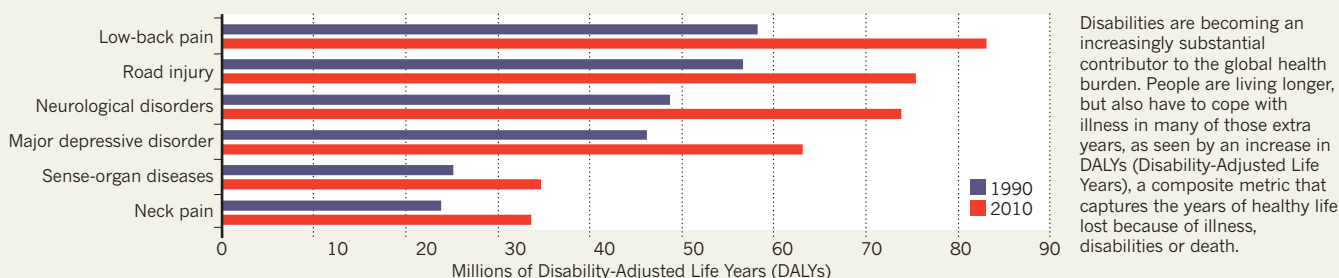


Life expectancies have risen in almost every region of the world, but have declined in countries including Haiti, Zimbabwe, Lesotho, Ukraine and Belarus.

— Indicates range of life expectancy in countries within the region.



With the death rate from infectious diseases on the decline, and people living longer, the world is in a 'disease transition'. Non-communicable diseases have supplanted infectious diseases, maternal mortality and infant diseases as the main causes of 'Years of Life Lost' — with the big exception being in sub-Saharan Africa.



Disabilities are becoming an increasingly substantial contributor to the global health burden. People are living longer, but also have to cope with illness in many of those extra years, as seen by an increase in DALYs (Disability-Adjusted Life Years), a composite metric that captures the years of healthy life lost because of illness, disabilities or death.

HEALTH

Global survey reveals impact of disability

Study tracks changes in life expectancy and health burdens.

BY DECLAN BUTLER

People around the world are generally living much longer than before — but many of those extra years are blighted by disability, according to the latest estimates of the global burden of disease, debilitating disorders and injuries.

The estimates, a major revision of a 1990 assessment, were published in *The Lancet* last week by an international collaboration involving 486 scientists from 302 institutions in 50 countries. The project was led by the Institute for Health Metrics and Evaluation at the University of Washington in Seattle (see go.nature.com/iem2sh).

In the 1970s, people in many developing countries could often hope to live only until their forties or fifties. Now they can generally expect the longer lifetimes once enjoyed only by those in the richest countries. The big exception is countries in sub-Saharan Africa, which have made much less progress (see 'A shifting picture of global health').

Factors such as economic growth, better education and improved health technologies have driven the gains by reducing deaths from infectious diseases, malnutrition and maternal and child illness. As people live longer, non-communicable diseases such as heart disease and cancer, and risk factors such as obesity and high blood pressure, have become

the major challenges facing health systems in much of the world.

The findings highlight another group of conditions, such as musculoskeletal ailments and mental illness — as well as injuries — that cause relatively fewer deaths but result in substantial disabilities. These conditions now deprive people of a similar number of years of healthy life to the classical killer non-communicable diseases. Living longer can be a net benefit for individuals, even if they are ill, but at the population level this translates into higher costs and major challenges for health-care systems.

Estimating the burden of death and illness remains an inexact science, however. This is mainly because of the lack of data — particularly from poorer countries that often do not have basic systems for certifying deaths and their causes — and wide disparities in the comparability and quality of data. Nonetheless, such estimates will undoubtedly help to shape health care, priorities for research and international aid, and funding allocations. Researchers and policy experts can thus be expected to pore over, debate and often contest the detailed findings in the coming months. ■ [SEE EDITORIAL P.311](#)

BUSINESS

China's patent boom brings legal wrangles

Court decision sets precedent for protecting intellectual property.

BY DAVID CYRANOSKI

China's ambitious efforts to build an 'innovation economy' have led to a surge in patents as companies seek to protect their hard-earned intellectual property (IP). From 2008 to 2011, the number of patent applications grew by an average of more than 20% each year; last year China handled 526,412 applications, overtaking the United States for the first time, according to a report issued last week by the World Intellectual Property Organization in Geneva, Switzerland (see 'Patent boom').

But that increase has been accompanied by a huge rise in patent litigation — 7,819 cases in 2011, according to China's Supreme People's Court, roughly twice as many as in the United States — as companies and researchers stumble through the unfamiliar terrain of IP law. Although one would expect a growing volume of patents to lead to more litigation, notes Elliot Papageorgiou, an IP-enforcement expert at law firm Rouse in Shanghai, Chinese companies are also "learning how to use IP as 'swords' against both local competitors as well as foreign companies".

A high-profile case settled last month highlights how companies court trouble by failing to protect their technologies, and how academic scientists can get dragged into the fray. The modest punishment meted out is also raising concerns that China's courts do not offer enough protection for innovators.

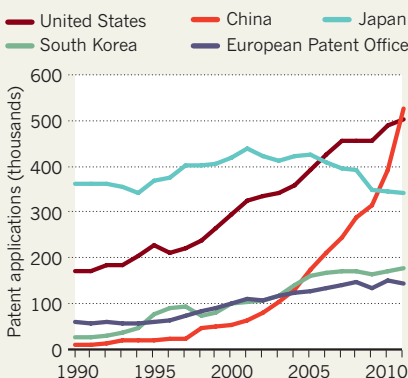
"China is well on the way to having a high-quality IP legal and enforcement structure," says Ian Harvey, co-director of Tsinghua Business School IP Centre in Beijing. "But some regions are further behind, with people and companies not understanding what IP is, how it is used in business and what they should be doing."

The recent case involved technology for producing a long-chain organic molecule called dodecanedioic acid (DC12), used to make nylon, lubricants and pharmaceuticals. In the late 1990s, Liu Xiucui, founder and chief executive of Cathay Industrial Biotech in Shanghai, developed a process to mass-produce the acid, winning major international customers such as the chemical company DuPont, based in Wilmington, Delaware.

NATURE.COM
Read more at [Nature China:
nature.com/nchina](http://NatureChina.nature.com/nchina)

PATENT BOOM

Patent applications have grown rapidly in China, and last year surpassed those in the United States.



Cathay now produces more than 10,000 tonnes of DC12 per year, accounting for about half of the world's supply. "Our process solved a problem that DuPont and others could not," says Liu. Cathay uses fermentation to convert fatty acids into dicarboxylic acids — and was able to scale up the procedure with the help of a particularly efficient strain of bacterium. But Cathay did not patent some key technologies in the process, instead choosing to keep them as trade secrets (an option that is not uncommon for companies around the world).

In 2010, Shandong Hilead Biotechnology in Laiyang announced that it would start producing 10,000 tonnes of dicarboxylic acids a year, with the intention of ramping up to 60,000. Hilead set its prices much lower than the market rate — forcing Cathay to do the same — and started to file patents on crucial parts of the fermentation process.

Liu suspected that Hilead had stolen his process, in part because Wang Zhizhou, a former Cathay employee, had resigned for personal reasons in 2008 and became Hilead's deputy general manager the following year. So Cathay filed a series of lawsuits to reclaim its intellectual property, embroiling not only various employees of both companies, but also the Chinese Academy of Sciences Institute of Microbiology (IMCAS) in Beijing.

A ruling in the first case has given Cathay a decisive victory. The case was brought by two Cathay employees, Lei Guang and Li Naiqiang, who claimed that a 2010 Hilead patent had

"denied their right, under Chinese patent law, to authorship of the processes they developed". Hilead's patent bore the names of multiple IMCAS staff members, including director Li Huang, and Hilead employees including Wang.

A July decision of the Beijing Number 1 Intermediate Court, upheld on 16 November by the Higher People's Court, found that the patent was "plagiarized" from Cathay's technical methods, and noted that Wang and Ge Minghua, another former Cathay employee who went to work for Hilead, had had access to the details of Cathay's procedures. Li Huang says that he did not even know that he had been named on the patent, and that IMCAS is taking steps to avoid similar situations in the future. None of the other defendants in the case replied to *Nature's* questions about their involvement.

Li Naiqiang says that the decision sets an important precedent for protecting the IP of researchers and companies in China: "If achievements are not protected, it will reduce industry willingness to do research and development and will hinder the development of the high-tech industry."

Cathay's victory is unlikely to make a big difference to its bottom line. Although the ruling invalidates Hilead's patent, preventing the company from using it to shut down Cathay's process, Hilead does not have to stop its own dicarboxylic acid production (Cathay hopes to achieve that through further legal action). Hilead is, however, required to apologize publicly in *Science and Technology Daily*, the official newspaper of the Ministry of Science and Technology. "In the Chinese mind, the moral punishment is measured in parallel with economic damages," says Wei Wei, a patent lawyer in Dallas, Texas, who handles cases related to China.

Zhiwu Chen, an expert on China's economy at Yale University in New Haven, Connecticut, agrees that the lack of monetary compensation is "pretty typical" in China. He says that this is partly because of the courts' reluctance to set precedents that could lead to penalties for state-owned enterprises.

But the country will achieve its potential for innovation only if the judicial system changes, Chen argues. If innovators are to feel secure in the country, he says, courts must start to calculate financial damages. "This example shows it doesn't pay to be an innovator in China." ■

366 DAYS: the year in science

BY RICHARD VAN NOORDEN

Two of the biggest breakthroughs of this leap year relied on breathtaking amounts of data. The ENCODE project has generated 15 terabytes of data over the past five years to uncover the functions of human DNA sequences; CERN has stored 26 petabytes of data this year alone from its Large Hadron Collider, as physicists worked to prove the existence of the Higgs boson. But data were a source of controversy as well as discovery. Arguments raged over whether information about a potentially dangerous flu virus should be published, for example, and funders, publishers and researchers discussed how to make raw data — as well as peer-reviewed research — more openly available. Meanwhile, high-profile cases of dubious or fraudulent results offered a reminder that above all else, findings need to be trustworthy.

THE HIGGS AT LAST Applause, relief, joy and tears: in July, the world's largest physics experiment officially discovered the Higgs boson. It took more than 500 trillion proton collisions at the Large Hadron Collider (LHC) at CERN, Europe's particle-physics lab near Geneva, before physicists could confidently announce that they had seen a new boson with a mass of around 125 gigaelectronvolts. Nearly 50 years ago, theorists including Peter Higgs had proposed that a Universe-filling quantum field imparts mass to some particles. The Higgs boson — the embodiment of that field — is looking disappointingly mundane so far, with no convincing hints of behaviour beyond that predicted by the standard model of particle physics. Nor has the LHC spotted evidence for the additional particles predicted by supersymmetry, a theory that would extend our understanding of the subatomic world and help to explain mysteries such as dark matter.

GOING TO EXTREMES In this Olympic year, science provided plenty of its own records. After two decades of drilling, a Russian team broke through 3.8 kilometres of Antarctic ice in February to reach Lake Vostok, a huge body of water isolated for millions of years. Early sample analysis has not found any signs of the life many scientists thought the lake might host. As *Nature* went to press, a British team hoping to reach Lake Ellsworth, one of the continent's other subglacial lakes, was battling technical problems with the high-pressure jet of hot water used to bore through the ice. Film director James Cameron, meanwhile, became the first person to dive solo to the deepest spot on the planet: the bottom of the Mariana Trench, almost 11 kilometres deep. Just as gripping — though less scientifically valuable — was skydiver Felix Baumgartner's jump from more than 39,000 metres above New Mexico, breaking the speed of sound and a height record held since 1960.

2012 in review



GRAHAM STUART/AFP/GETTY

Vindicated: Peter Higgs's prediction gained weight this year.

But not every record-beating attempt was successful. After six years of trying, the US\$3.5-billion US National Ignition Facility in California — the world's most powerful laser — failed to meet its target of achieving 'ignition', a fusion power milestone in which a small pellet of hydrogen isotopes blasted by the laser would generate as much fusion energy as the beams put in.

RETHINKING ENERGY Nations' energy policies are continuing to shift in the wake of last year's nuclear disaster in Fukushima, with Japan outlining options for a future almost free of nuclear energy. The country switched off its last operating nuclear reactor for maintenance work in May, and faced widespread public protests against turning any reactors back on — although it did manage to restart two in July. In Europe, stress tests of more than 140 reactors concluded that widespread safety

172 ▶ NUMBER OF PAPERS IN WHICH YOSHITAKA FUJII IS THOUGHT TO HAVE FABRICATED RESULTS

392.9 ▶ ATMOSPHERIC CONCENTRATION OF CARBON DIOXIDE (IN PARTS PER MILLION) — A RECORD HIGH



Superstorm Sandy ravaged New York in October.

upgrades are needed. The US Nuclear Regulatory Commission, meanwhile, granted a licence for a plant that uses lasers to enrich uranium for nuclear fuel, a technology that some fear could enable bomb-makers to covertly enrich uranium. Countries also continued to explore unconventional sources of gas and oil to keep the lights burning and cars on the road. The United States proposed rules for the booming shale-gas fracking industry, which has enabled the US electric-power industry to shift 10% of its generating capacity from coal to gas. According to the International Energy Agency, the United States is also on course to be the world's largest oil producer by 2020, and almost self-sufficient in energy by 2035. But there were reminders of the dangers of searching for new oil reserves. Shell was unable to begin its drilling programme in the Arctic sea after damage to drilling vessels, and BP was hit by US\$4 billion in criminal fines relating to the April 2010 Deepwater Horizon

oil accident in the Gulf of Mexico. The fledgling clean-energy industry had its own problems: lithium-battery maker A123 Systems of Waltham, Massachusetts, went bankrupt in October as the market for electric cars remained small.

Curiosity's arrival on Mars was cause for celebration at NASA.

DATA ON DISPLAY Science, famously, self-corrects. By March, researchers had firmly scotched last year's suggestion that neutrinos might travel faster than light, and a number of experiments had refuted the 2010 claim that a bacterium can use arsenic in its DNA. But correction does not always come so quickly: when studies are hard to replicate, bias or error can linger for years. Anaesthesiologist Yoshitaka Fujii and nutrition researcher Eric Smart were both censured this year for decades of misconduct that had gone undetected until relatively recently, and psychology came in for particular criticism after a number of accusations of massaged data led to some high-profile resignations. Scientists worried more generally about the issue of irreproducible results and set up efforts such as the Reproducibility Initiative to get independent labs to replicate high-profile research. The idea that scientists should communicate and publish their data more openly also gained momentum. High-profile online open-access journals such as *eLife* and *PeerJ* launched, and the open-access movement made headway in Britain, where government and private research funders said in July that they would pay for papers to be made publicly accessible from April 2013.

THE ROVER HAS LANDED "It's the wheel! It's the wheel!" came the cry, as NASA scientists saw the first images of their Curiosity rover on the surface of Mars, where a hovering sky crane had gently deposited it. Since landing at Gale Crater in August, Curiosity has provided startling images and analysis of the Martian surface and atmosphere, but has not yet found any methane or organic molecules that might hint at

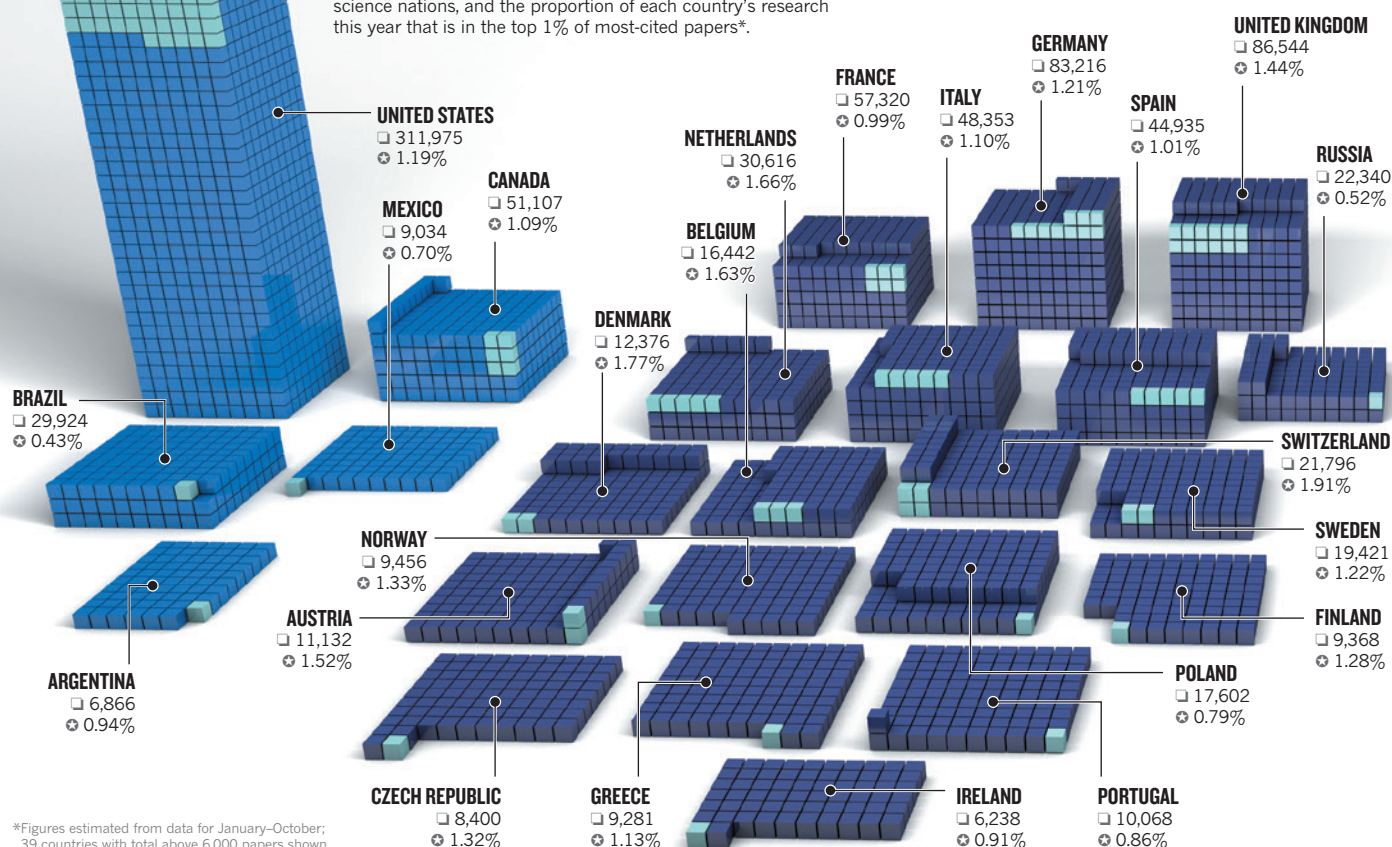
NATURE.COM
For an interactive
look at the year in
numbers:
go.nature.com/xrggmf

1,640 ► NUMBER OF
GENOME-WIDE
DATA SETS IN
ENCODE

2,932 ► NUMBER OF AUTHORS ON THE
ATLAS COLLABORATION PAPER
THAT ANNOUNCED THE
DISCOVERY OF THE HIGGS

SCIENTIFIC PAPER TRAIL

Number of research papers published in 2012 by leading science nations, and the proportion of each country's research this year that is in the top 1% of most-cited papers*.



the presence of life. Outside the Solar System, new planets swam into focus thanks to the eagle-eye of the Kepler space telescope, which has now racked up some 3,000 candidate new worlds. Earth-based scopes played a major part, too, discovering an Earth-sized exoplanet in our neighbouring star system, α Centauri, a mere 1.34 parsecs (4.4 light years) away. Among space successes, NASA's Dawn craft found evidence of water on the asteroid Vesta, and China sent its first female astronaut, Liu Yang, into orbit. But it was a private firm, not a government, that made the year's headline launch: in October, SpaceX sent its capsule Dragon on the first commercial resupply mission to the International Space Station. The craft has been so successful that the company, based in Hawthorne, California, is even considering a trip to Mars.

THE EARTH UNDER PRESSURE The summer's dramatic melt of sea ice in the Arctic set a record that exceeded the predictions of climate models, and the United States faced its most extensive drought in half a century. But for many east-coast Americans it was Superstorm Sandy, in late October, that came to embody the nebulous threat of global warming. The storm, which caused \$50 billion in damage, triggered discussion that focused more on how to adapt to increasingly likely weather extremes than how to prevent climate change. Global talks on the environment were overshadowed by economic concerns. At June's United Nations Earth Summit in Rio, developing countries argued against constraining their growth, and rich countries were reluctant to pledge more development aid. Similar political inertia hit climate talks in Doha in December, although delegates did agree to extend a weakened Kyoto Protocol to 2020. Individual countries did more: Mexico set legally binding emissions cuts in April, for example. In Brazil, the rate of deforestation in the Amazon fell to another record low this year, although in October the country passed

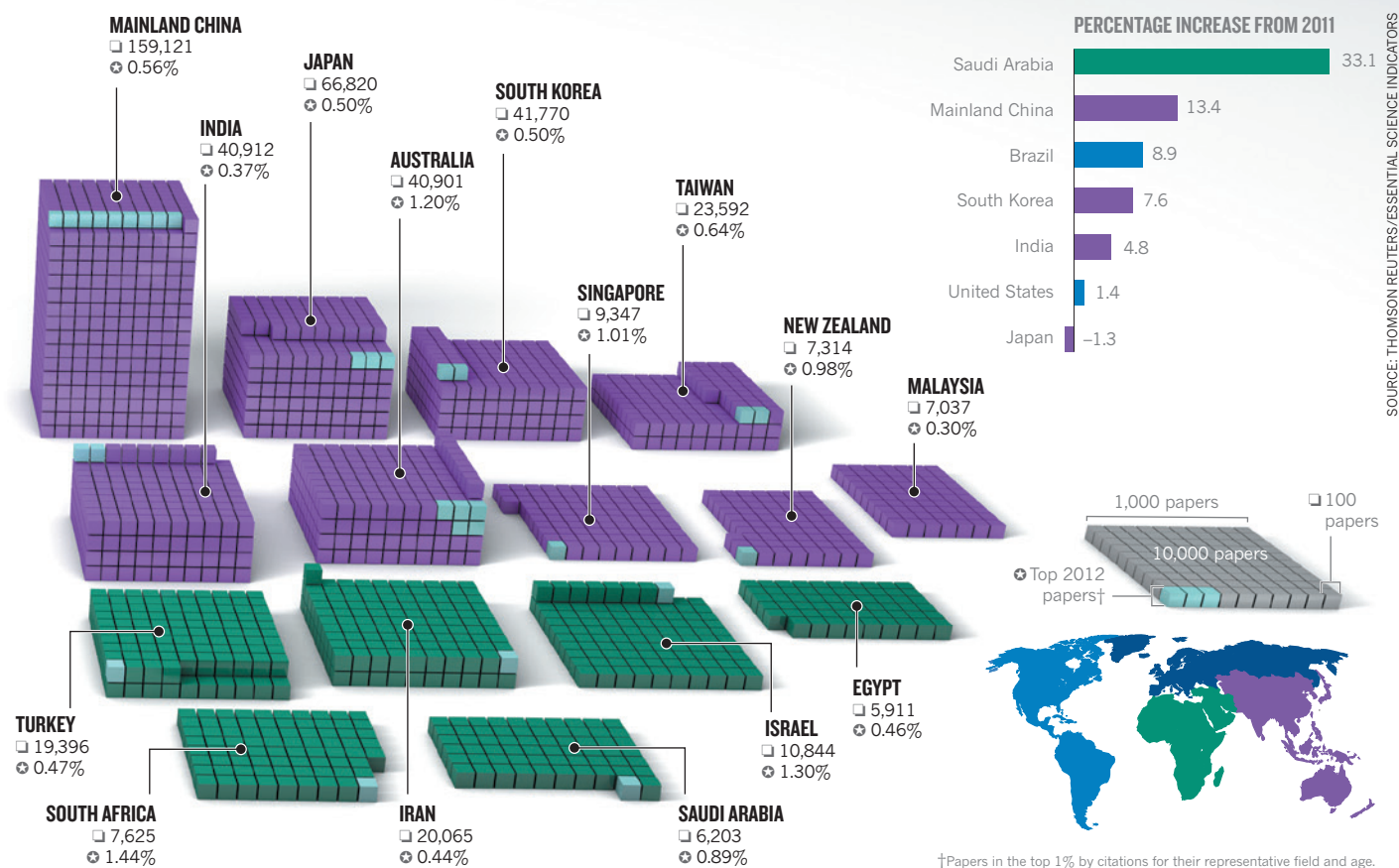
Lonesome George died in June.



a controversial act that may weaken protection for forests. In June, Australia unveiled plans for the world's largest network of marine reserves — but proposals for international protection of three large areas in the waters around Antarctica were knocked back in November. And in the Galapagos Islands, the death in June of the iconic giant tortoise Lonesome George — the last of his subspecies — called attention to the plight of endangered species around the world.

SCIENTISTS SPEAK OUT Many researchers prefer to keep their heads down when scientific controversies blow up, but they can certainly put up a fight when whole fields or scientific values are threatened. In May, UK scientists spoke up for the value of their work on genetically modified (GM) wheat when an anti-GM campaign group, Take The Flour Back, threatened to destroy it. And in October, researchers across the world reacted with dismay when an Italian court sentenced a group of experts to jail for six years, for allegedly playing down seismic risks before the devastating earthquake in L'Aquila in 2009. Critics warned that the precedent might make scientists reluctant to offer expert opinions for fear of prosecution or reprisals. But scientists have been quieter about other challenges: slowly but surely, animal-rights activists have this year restricted the transport of lab animals by a large number of cargo carriers, without facing effective opposition.

THE HUMAN ENCYCLOPAEDIA Little more than 1% of the human genome's 3 billion letters of DNA — just 20,000 genes — code for proteins. But vast regions of non-coding sequences still have a vital function, affecting the way the genome is packaged, regulated and read in different cell types. In September, a consortium of some 440 scientists released 30 papers from the ENCODE project (the Encyclopedia of DNA Elements), estimating that at least 20% of the genome can influence gene expression. Other ambitious projects to crunch big biological data included the first results from an effort to map the circuit wiring of the entire mouse brain, and a project to



SOURCE: THOMSON REUTERS/ESSENTIAL SCIENCE INDICATORS

track gene activity in some 900 anatomical parts of the human brain. At the cellular level, the flexibility of stem cells continued to astonish. US researchers found stem cells in women's ovaries that seem to be able to produce new eggs, contradicting the dogma that women are born with their life's supply of gametes. And Japanese scientists showed how to coax stem cells from mice into becoming viable eggs. Fertilized and transplanted back into foster mother mice, they produced healthy offspring.

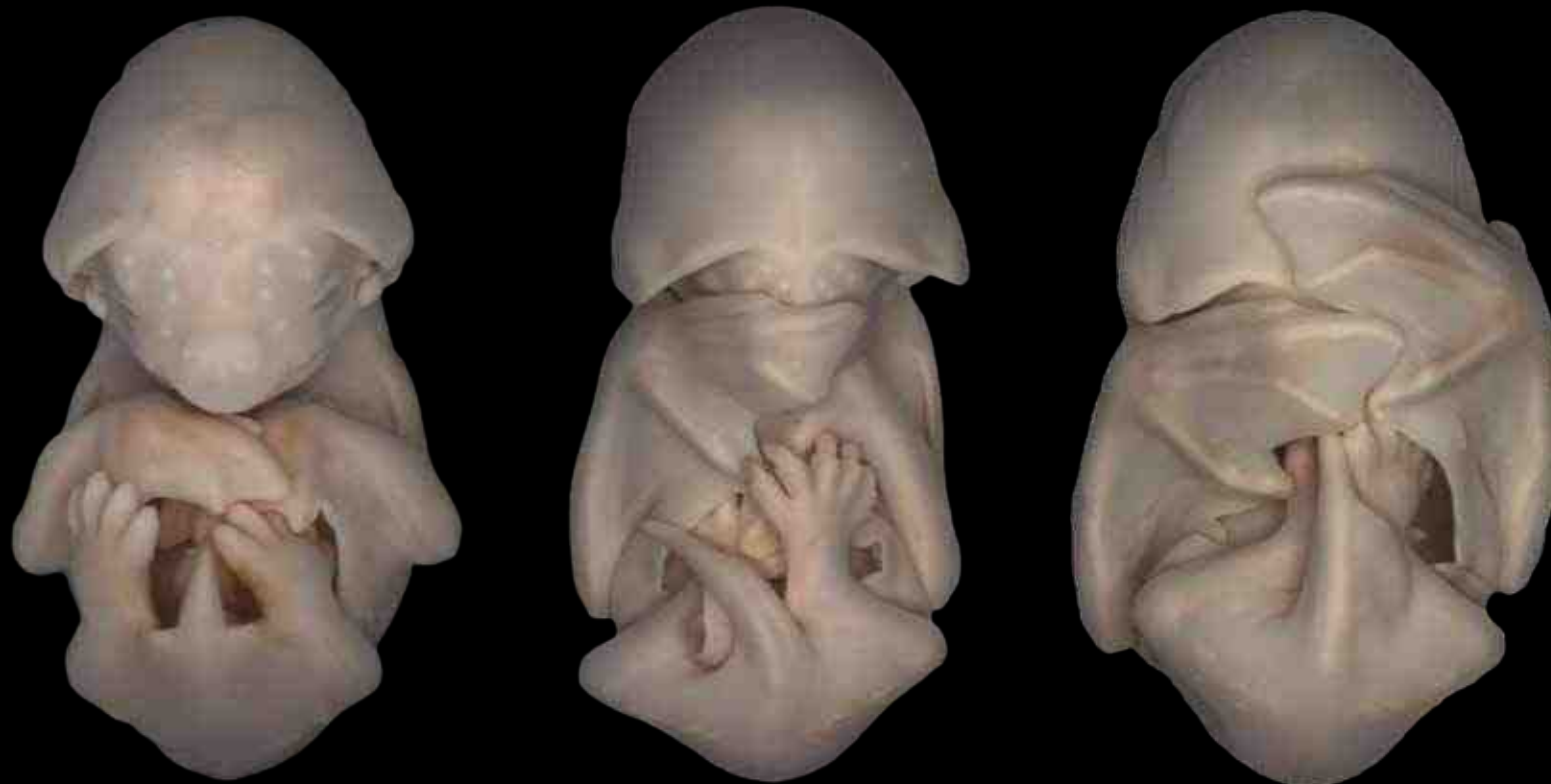
VIRAL STRIFE Two papers showing how mutated versions of the highly pathogenic H5N1 avian influenza virus can transmit between ferrets sparked international strife and a bitter intragovernmental feud in the United States. Some feared that disseminating the recipe for a mammalian-transmissible H5N1 would aid terrorists or increase the likelihood of an accidental release. And at the end of 2011, the US National Science Advisory Board for Biosecurity (NSABB) recommended that the papers be published only in redacted form. But others said that censoring the studies would fly in the face of the scientific ideal of open communication, and shut down potentially life-saving research. The NSABB reversed its position in March, and the papers were published in May and June. But the controversy continued: politicians lambasted the US government for acting too hastily in approving publication, while some scientists chided it for taking too long to reach a final decision. Government regulators are now considering tightening the restrictions for work on such viruses. While the details are being hashed out, a 'voluntary' moratorium on similar research has been in effect since January, angering some scientists who are anxious to get back to work.

AGE OF AUSTERITY As rich nations scaled back their public spending, research funding was also cut — although not in every country. Canada slashed spending on the environment and shut down a string of research programmes, including the renowned Experimental Lakes Area, a collection of 58 remote freshwater lakes in Ontario used to study pollutants for more than 40 years. Spain's 2013 budget proposal would reduce research funds for a fourth consecutive year, and follows a 25%

cut in 2012. In the United States, scientists spent most of the year worrying about the 'sequester', an across-the-board budget cut that may take effect early next year, although some cuts were made in this election year too: NASA's planetary scientists held a cake sale to highlight their field's dwindling support. Talks on the enormous 2014–20 European budget — including a proposed €80 billion (US\$104 billion) for research under the Horizon 2020 programme — broke down in November and will restart in 2013. Even India scaled down its historic funding growth to more cautious inflation-level increases for 2012–13. But it wasn't all bad news: China's central government boosted its spending on science by nearly 12.5%; France's 2013 austerity budget still found room to boost science cash by 2.2%; Germany channelled more federal funding to universities (creating a large health-sciences institute in Berlin); and the US biotech sector saw the glimmers of a revival from public-market investors.

PHARMA'S FUTURE The US Food and Drug Administration approved two weight-loss drugs — Belviq (lorcaserin) and Qsymia (phentermine plus topiramate) — this year, the first since 1999. The agency also gave a green light to Truvada, the first drug designed to prevent HIV infection. But two monoclonal antibodies designed to fight Alzheimer's disease, bapineuzumab and solanezumab, failed keenly awaited clinical trials — although solanezumab may have slowed cognitive decline in some cases. Researchers think that preventing Alzheimer's at an earlier stage could be a more promising strategy, and hope to set up pre-emptive trials in 2013. Among significant business moves, California-based sequencing company Complete Genomics went to China's BGI for \$118 million, despite competition from Illumina; biotech giant Amgen said that it would buy deCODE Genetics for \$415 million; Bristol-Myers Squibb and AstraZeneca paid US\$5.3 billion to acquire biotech firm Amylin; and GlaxoSmithKline got Human Genome Sciences in a deal worth US\$3.6 billion. Pharmaceutical companies also paid a record amount in malpractice fines in the United States this year. ■

Additional reporting by Brendan Maher



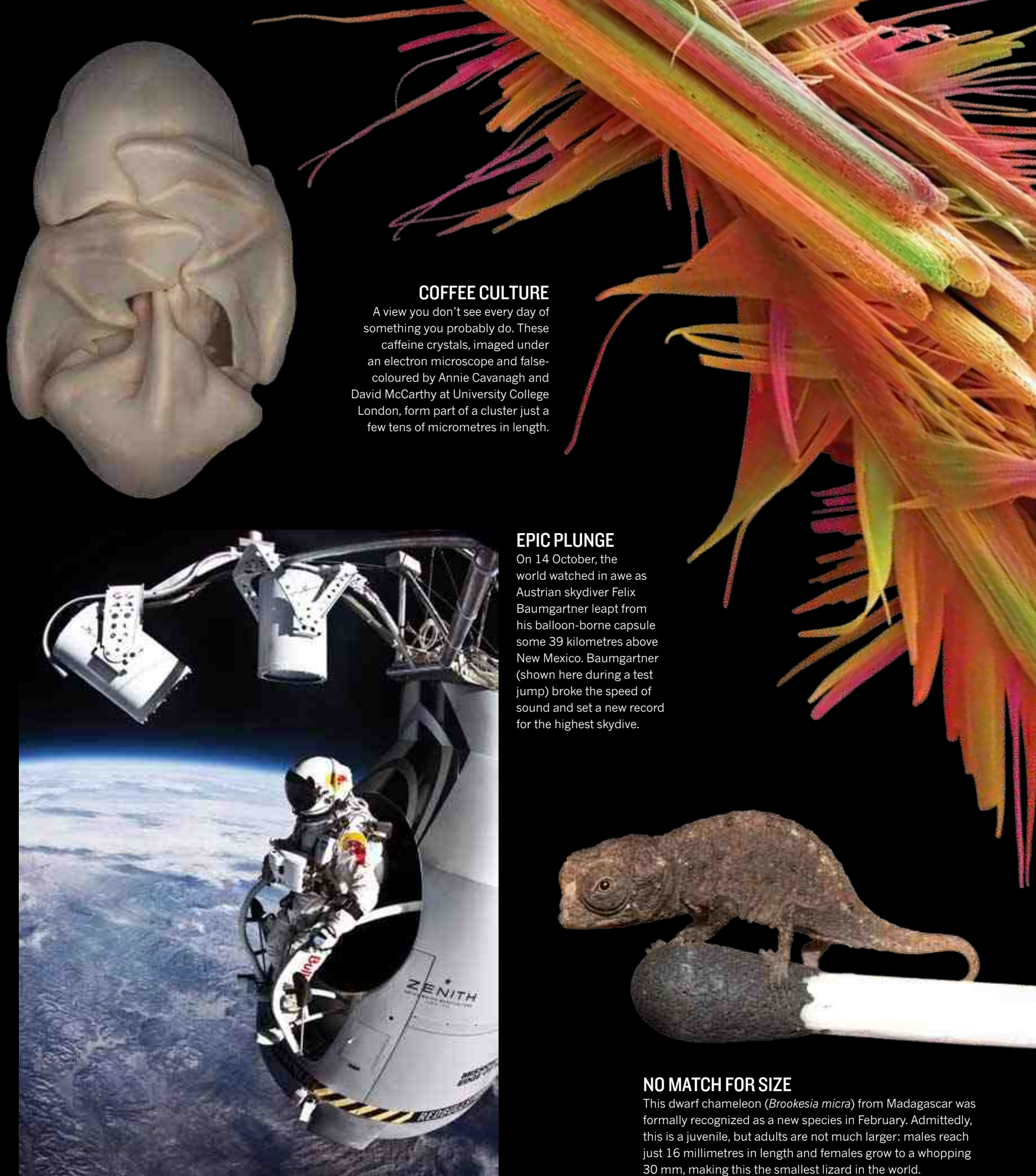
FROM THE WOMBS OF BATS

These apparitions — seemingly dreamed up by Swiss surrealist H. R. Giger — are embryos of the black mastiff bat (*Molossus rufus*). Biologist Dorit Hockman of the University of Cambridge, UK, took the picture using a standard dissecting microscope for inclusion in a system that documents embryo development.



THUNDER GOD

This image of the Thor's Helmet Nebula (NGC 2359) was taken by the Very Large Telescope array on Mount Paranal in Chile, on the occasion of the 50th anniversary of the European Southern Observatory this year. The organization's next major project — the Atacama Large Millimeter Array — is due to be completed in 2013.



COFFEE CULTURE

A view you don't see every day of something you probably do. These caffeine crystals, imaged under an electron microscope and false-coloured by Annie Cavanagh and David McCarthy at University College London, form part of a cluster just a few tens of micrometres in length.

EPIC PLUNGE

On 14 October, the world watched in awe as Austrian skydiver Felix Baumgartner leapt from his balloon-borne capsule some 39 kilometres above New Mexico. Baumgartner (shown here during a test jump) broke the speed of sound and set a new record for the highest skydive.

NO MATCH FOR SIZE

This dwarf chameleon (*Brookesia micra*) from Madagascar was formally recognized as a new species in February. Admittedly, this is a juvenile, but adults are not much larger: males reach just 16 millimetres in length and females grow to a whopping 30 mm, making this the smallest lizard in the world.



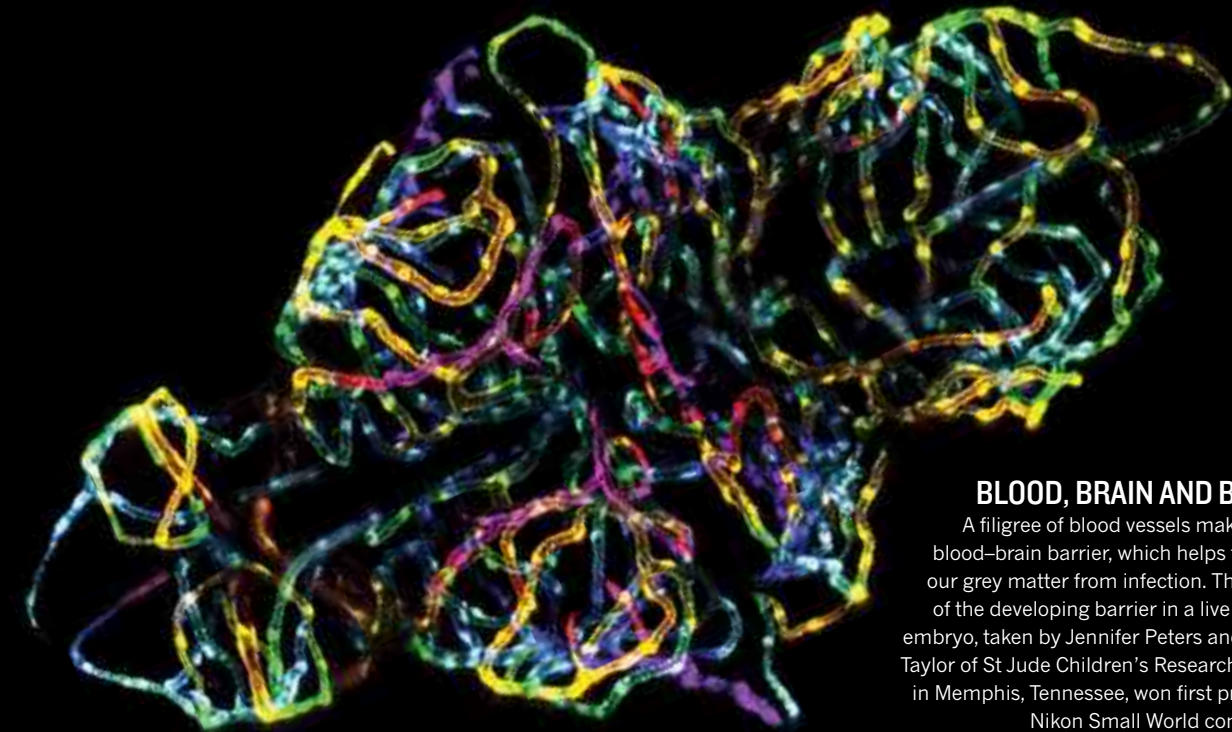
FREEZE FRAME

The extent of Arctic sea ice hit a record low this year, pegged at 3.41 million square kilometres on 16 September by the US National Snow and Ice Data Center in Boulder, Colorado. This is 3.29 million square kilometres less than the average extent recorded between 1979 and 2000.

IMAGES OF THE YEAR

With every moment of our lives seemingly Instagram'd, Facebooked and Twitpic'd, has the power of the photographic image faded like a Polaroid ravaged by sunlight? Not a bit of it. The camera is always on hand to record the wonder of the natural world and the thrill of exploring it. This year it showed us a man falling to Earth faster than the speed of sound; the nightmarish denizens of the deep oceans; and the delicate tracery of a brain's protective shield. These, and more, are some of *Nature's* favourite pictures of 2012.

Images selected by *Nature's* art and design team
Text by Daniel Cressey and Mark Peplow



BLOOD, BRAIN AND BEAUTY

A filigree of blood vessels makes up the blood-brain barrier, which helps to protect our grey matter from infection. This picture of the developing barrier in a live zebrafish embryo, taken by Jennifer Peters and Michael Taylor of St Jude Children's Research Hospital in Memphis, Tennessee, won first prize in the Nikon Small World competition.



NOT IN KANSAS ANY MORE

Nebraska, actually. The United States saw its share of extreme weather this year, from the most extensive drought in half a century to the violence of Superstorm Sandy. Fortunately, this tornado-spawning mesocyclone in June passed farm buildings in Gurley without giving them a one-way trip to Oz.

PLASMA BURST

This solar filament, some 350,000 kilometres long, erupted from the surface of the Sun on 31 August. Seen in the extreme ultraviolet by NASA's Solar Dynamics Observatory satellite, the eruption became a coronal mass ejection moving at about 1,400 kilometres per second — its particles grazed Earth's magnetosphere several days later, sparking an auroral display.



AND THIS IS ME ON MARS

Since it landed on 6 August after its '7 minutes of terror' descent, NASA's Mars rover Curiosity has delighted scientists and the public with streams of data and images from the red planet.



MONSTER RAGWORM

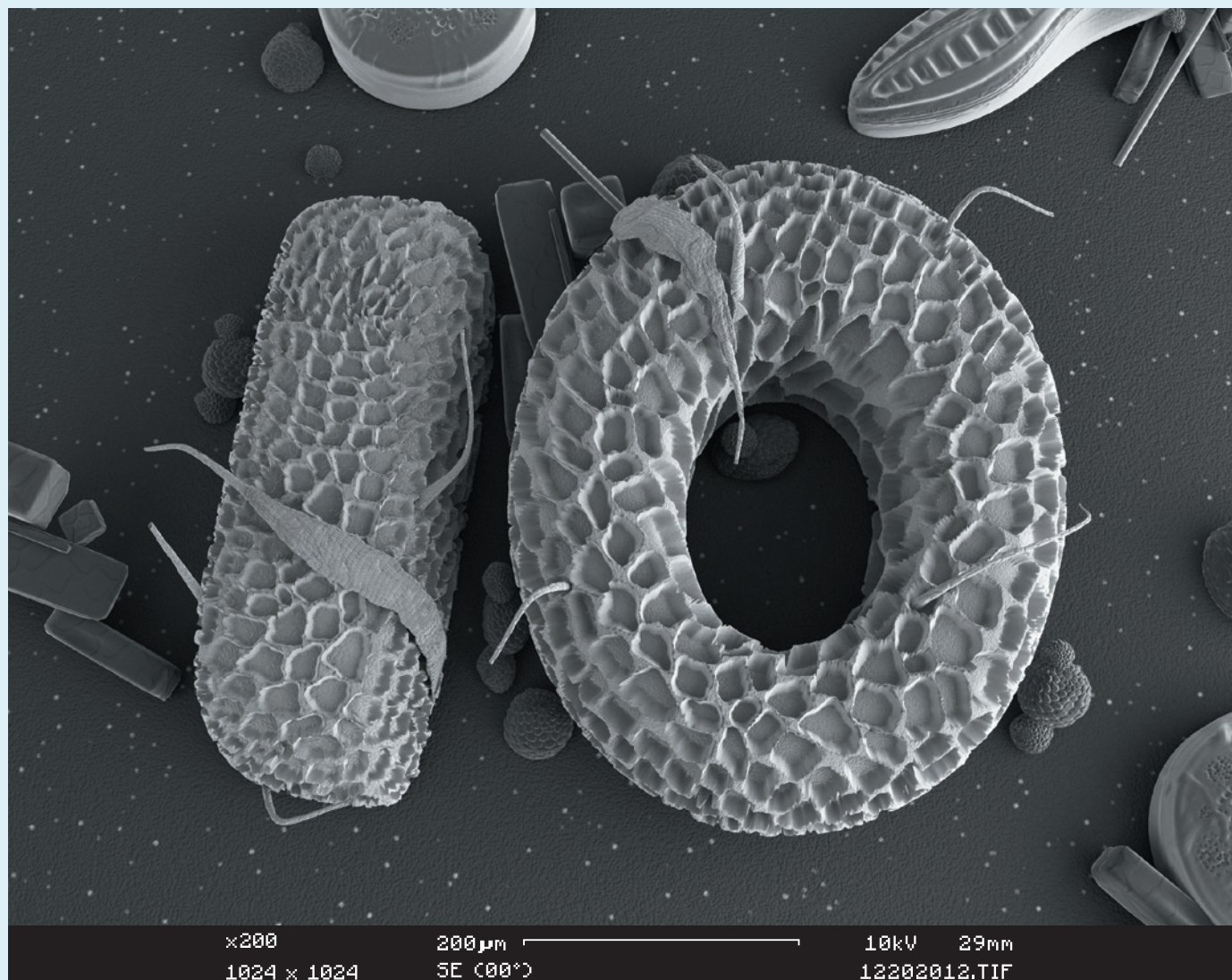
This specimen of *Nereis pelagica* was found in the White Sea off the Russian coast. Photographer Alexander Semenov says that the predatory worm was about 25 cm long and as thick as a finger. It is not known whether he still has nightmares about it.

ARCTIC ICE: ANNA HENLY; WINNER, THE WORLD IN OUR HANDS AWARD/VEOLIA ENVIRONNEMENT WILDLIFE PHOTOGRAPHER OF THE YEAR 2012/BBC WORLDWIDE/NATURAL HISTORY MUSEUM; BATS: DORIT HOCKMAN/UNIV. CAMBRIDGE/2012 PHOTOMICROGRAPHY COMPETITION/NIKON SMALL WORLD; NEBULA: B. BAILLEUL/ESO; BAUMGARTNER: JAY NEMETH/RED BULL CONTENT POOL.

CAFFEINE: ANNE CAVANAGH/DAVID MCCARTHY/HELLCOME IMAGES AWARDS 2012/HELLCOVE IMAGES; CHAMELEON: FRANK GLAW; BLOOD-BRAIN BARRIER: JENNIFER L. PETERS/MICHAEL R. TAYLOR/ST. JUDE CHILDREN'S RESEARCH HOSPITAL/NIKON SMALL WORLD; CYCLONE: CAMILLE SEAMAN; SUN: SDO, AIA TEAM/GSFC/NASA; CURIOSITY: MALIN SPACE SCIENCE SYSTEMS/JPL-CALTECH/NASA; RAGWORM: ALEXANDER SEMENOV

ROLF-DIETER HEUER / CYNTHIA ROSENZWEIG / ADAM STELTZNER / CÉDRIC BLANPAIN / ELIZABETH IORNS
JUN WANG / JO HANDELSMAN / TIM GOWERS / BERNARDO DE BERNARDINIS / RON FOUCHIER

CHRIS LABROOY



NATURE'S 10

Ten people who mattered this year.

366 DAYS:
the year in science



**ROLF-DIETER
HEUER**

LAURENT EGLI/CERN

THE HIGGS DIPLOMAT

Gentle nudging from the head of CERN ensured that the world heard about the discovery of a long-sought particle.

BY GEOFF BRUMFIEL

The detection of the Higgs boson, announced to the world on 4 July, was the scientific discovery of the year, if not the decade. But it might not have been a discovery at all had it not been for the diplomacy of Rolf-Dieter Heuer.

As the director-general of CERN, the particle-physics laboratory located near Geneva, Switzerland, Heuer has budgetary control over the Large Hadron Collider (LHC), which generated the Higgs by smashing together protons at energies higher than those of any other accelerator in the world. But Heuer has considerably less authority over the two experiments that detected the particle in the debris of the collision; those facilities are run democratically by the thousands of physicists who built them.

Heuer didn't even know what the groups had seen until mid-June, when their two sleep-deprived leaders, Fabiola Gianotti and Joe Incandela, gave him short presentations on what their building-sized detectors had found. The results were tantalizing: after months of gathering data, both experiments now had strong signals from a new kind of particle inside the LHC — signals that closely mirrored what theorists expected from the Higgs, first predicted nearly 50 years ago.

Yet neither group leader was willing to claim that they had 'discovered' the Higgs. High-energy physicists are reluctant to make such a declaration until the chances of a statistical error are whittled down to a level known as 5-sigma — 1 part in 3.5 million — and neither experiment was quite there with the Higgs data. Both leaders had many group members pushing to delay announcement of the discovery until the end of the year, when the steadily accumulating data would make the

conclusion cast iron. "Fabiola and I felt tremendously stressed about the whole thing," Incandela recalls.

Heuer was under pressure, too. That same week, there was a meeting of CERN's governing council: a committee of civil servants, diplomats and independent scientists from the lab's member states. The council made it clear that, if there was something to say about the Higgs, the physicists had a duty to say it out loud to the governments and citizens who had paid for the LHC.

Heuer and the detector groups agreed to have a public seminar at CERN on 4 July, but with days to go it was still unclear just what would be claimed. The evidence was growing stronger every day as more data were analysed, but the detector groups remained cautious and reluctant to make any bold claims. A less astute director-general might have pushed Incandela and Gianotti to declare victory, says Incandela. "Imagine, you've got these two people coming in who are both paranoid as hell and incredibly conscientious showing you these unbelievable results," he says. "Rolf was looking at us both and thinking, 'What's the matter with these guys?'"

But the affable, 64-year-old Heuer has a reputation for guiding without being bossy, and for listening. "In German it's *'ein Händchen'*" — a gentle hand", he says. "You have to have a feeling for people, a feeling for what they can do and what they cannot do." If the consensus within the experiments wasn't quite there, he wasn't going to push it further. But after consulting with Incandela, Gianotti and others, says Heuer, "I decided that I could use the word 'discovery'." He would take the risk that the experiment groups wouldn't.

The final press release was drafted with just two days to go and, as promised, the word "discovery" appeared just once, in a quote from Heuer. On the morning of the announcement, Gianotti and Incandela stood before their peers, politicians and the press. In back-to-back talks, they laid out the evidence they had for the new particle. "We agreed that we would just stick to the facts," says Gianotti.

After the talks, Heuer, who had played the role of jovial master of ceremonies for the entire affair, stood before the packed auditorium. Now was the time for him to drop the 'D' word. But he paused ahead of his prepared remarks. Ever the diplomat, his first words were not a declaration but a question: "As a layman, I would now say, 'I think we have it.' Do you agree?"

The auditorium burst into applause. ■

GUARDIAN OF GOTHAM

New York's climate-adaptation champion is determined to make her city more resilient to natural disasters.

BY JEFF TOLLEFSON

As Superstorm Sandy battered the US east coast this year, Cynthia Rosenzweig huddled with her 97-year-old mother in a suburb of New York City, not far from where she grew up. After making sure that her own home had sustained only minor damage, Rosenzweig turned her attention to the city, which had not been so lucky.

Sandy had driven a 4-metre wall of water into low-lying neighbourhoods, destroying homes, flooding transportation tunnels and leaving millions of people without power. Although the damage came as a shock to most, Rosenzweig and a team of researchers had forecast those consequences a dozen years earlier as part of the first national assessment by the US Global Change Research Program.

"Everything that happened is in our earliest report," says Rosenzweig. Because of that work and many follow-on studies conducted for state and city officials, New York has incorporated climate-change adaptation and resilience into its long-term planning initiatives, which include upgrading building codes and managing parks and wetlands to accommodate flooding and sea-level rise. The actions have made New York a leader among cities working to prepare for the threats of climate change, says Rosenzweig. She is now trying to assess whether these steps helped to lessen Sandy's impacts, which may offer a preview of the threats expected as climate change intensifies storms and raises sea levels.

Rosenzweig's path to urban protector started in the fields of Tuscany, Italy, in 1969. She had left university in California to rent a small farm with her future husband, where they learned to pick grapes and olives and raise goats, pigs, ducks and geese. Eventually, she decided it was time to go back to

university to study agriculture. While pursuing a master's degree at Rutgers University in New Jersey, she found her way to a job at NASA's Goddard Institute for Space Studies (GISS) in New York, analysing satellite data on croplands.

When she arrived, GISS director James Hansen was busy modelling the impacts of a doubling in atmospheric carbon dioxide concentrations. He wanted to know what would happen to crops. As an agriculture expert, "I was the only person at GISS at that time who could begin to answer that question", she says. "I've been answering that question ever since."

To do so, Rosenzweig had to expand her focus from crops to agricultural economics, folding in the broader impacts on farmers, food supply systems and society. Thanks to her experience in complex assessments, Rosenzweig was chosen to head up the northeastern regional analysis when the first national climate assessment kicked off in 1997. Her team's analysis suggested that singular shocks such as Sandy would cause widespread problems.

Today, she co-chairs the New York City Panel on Climate Change, which advises local policy-makers. She is also helping to coordinate regional and international groups that are exploring climate adaptation and resilience.

The city has a long way to go. Rather than focusing on big-ticket solutions such as storm-surge barriers, Rosenzweig calls for a range of initiatives, from increasing redundancy in the electric grid to sealing off tunnels and making coastal areas more resilient to flooding. She doesn't pause for breath while running through the litany of needs. "We have to do more. We have to do better. We have to spend more money. We need pilot funding projects. There's just a lot to it." Then Rosenzweig flashes a wide smile. "I'm not a pessimistic person," she says. "We have to succeed. We don't have a choice." ■

NATL. CLIMATE CHANGE ADAPTATION FACILITY





OUR MAN ON MARS

NASA called on an unconventional engineer to give its pricey new rover a baby-soft landing on the red planet.

BY ERIC HAND

When NASA's Curiosity rover slammed into the Martian atmosphere at 5,900 metres per second on the night of 5 August, no one on Earth was more worried than Adam Steltzner, the engineer who led the 50-person entry, descent and landing (EDL) team for the US\$2.5-billion mission.

BILL INGALLS/NASA

At first, says Steltzner, all the data coming back from the spacecraft seemed fine as it began to brake by carving S-turns in the thin air like a crazed snowboarder. But then a few abnormal telemetry signals arrived at the control room at the Jet Propulsion Laboratory in Pasadena, California. "Beta out of bounds catastrophic," read one. Translation: Curiosity was tilting too much to the side. "The first two data points we see are like: 'You're going to die,'" says Steltzner.

He still uses the present tense as he recollects the landing sequence, and is not averse to hyperbole and dramatic pauses, unlike most engineers, who cloak their emotions in cautious jargon. "He's a very instinctual person," says colleague Al Chen, who was the voice of mission control that night. "And his instincts are often right."

Part of Steltzner's persona stems from growing up in Marin County, north of San Francisco. More interested in theatre and rock music than in his studies, the feckless teenager barely graduated from high school. Eventually, he realized he was just "bright and bored". He went

CELL TRACKER

By tracing the descendants of a single cell, a cautious developmental biologist tackled a controversy over tumour growth.

BY MONYA BAKER

Cédric Blanpain likes to see things for himself. His students, he says, "tell me in the morning that they have something very important. And I run, not to the screen, but to the microscope." When he was first setting up his lab at the Free University of Brussels six years ago, Blanpain's see-it-to-believe-it approach drew him to study untouched cells in tissues from living animals rather than, as is often done, in dishes or after transplantation. He didn't trust what cells do outside their normal environment.

His advisers warned him that the task would be tough — but Blanpain, a quick-talking developmental biologist with a penchant for snowboarding and jazz, was undeterred. He decided to refine a technique called lineage tracing, which reveals patterns of cell division in tissue. Blanpain uses low levels of a drug to activate a gene and change the colour of

specific cells and all their descendants, so that they can be seen under a microscope. He often works with a theoretical physicist to analyse the starting cells' contribution to the resulting tissue. No one has been able to track cell lineages as carefully or as quantitatively, says Brigid Hogan, a cell biologist at Duke University in Durham, North Carolina.

In October last year, Blanpain's team looked at mammary glands in mice during fetal development, pregnancy and lactation. Cell-culture work had led researchers to believe that a common cell type gives rise to multiple kinds of mammary cells. But when Blanpain looked at cells left undisturbed in mammary fat pads, he saw that they contributed just to a single lineage. He then went on to show that the adult mammary gland actually contains distinct types of stem cells (A. Van Keymeulen *et al. Nature* **479**, 189–193; 2011), a fact that could help to pin down the genesis of breast cancers. Blanpain "cleared up what had been a very confusing field", says Hogan.

This year, Blanpain tackled a long-standing controversy over the existence of cancer stem cells. By applying a carcinogen to mouse skin and then using his cell-tracking method, his team was able to show that cells do not contribute equally to the resultant tumours: some of the cells in a tumour peter out after a few divisions, and others — the stem cells — produce thousands of clones (G. Driessens *et al. Nature* **488**, 527–530; 2012). This implies that drug developers should focus on killing these tumour-generating cells.

Blanpain says that he hadn't expected such dramatic results. "I saw the first slide, and I said 'show me the second one'. After the fifth, I was sure what I was seeing." ■

on to earn a PhD in engineering physics and deeply absorbed the Navier–Stokes equations on fluid motion and the laws of thermodynamics until, he says, “they were woven into my soul”. But he still hasn’t given up his flair for the dramatic. Nor, with his western-style shirts, pomaded hair and big-buckle belts, has he ceded a rock star’s fashion sense. He describes his look as “rockabilly meets post-punk Poindexter”.

On the night of the landing, however, Steltzner was wearing one of the several hundred blue Curiosity golf shirts that had been issued to the mission team. Steltzner had had the acronym ‘EDL’ discreetly embroidered on the left sleeve of his team’s shirts — a badge of honour for the tightly knit group.

Steltzner paced around the control room. The dire tilt warnings turned out to be false alarms. But a second moment of tension came 7 minutes later, as a hovering jet pack called a sky crane slowly unspooled the 1-tonne rover to the floor of Gale crater from an altitude of 20 metres.

Nothing like the sky crane had been tried before in a Mars landing. Steltzner vividly recalls the brainstorming session in which he and a dozen others “groped it into existence”. In guiding other design choices, Steltzner tended towards a ruthless simplicity; he argued for the spacecraft to use one parachute instead of two, and for the rover to use a pulverizing drill instead of one that extracts a core. But with the sky crane, Steltzner can’t help but acknowledge a design sensibility that was a bit more baroque. “Because it looked so outlandish, we all felt very exposed,” he says. “If it failed, people would have been like, ‘You idiots.’”

That’s one reason why, in the control room, Steltzner wouldn’t declare success until he had two independent confirmations that the rover had landed, one from the rover itself and one from the sky crane, still hovering overhead. Then came the hardest part: he counted to ten, to ensure that the sky crane, after being severed and rocketed away, hadn’t somehow crashed down on top of the rover. He recalls, “At that point, I pointed at Al and said, ‘Do it.’ And he called out: ‘Touchdown confirmed. We are safe on Mars.’” ■

ELIZABETH IORNS

REPLICATION HOUND

A geneticist takes her quest to check results to the forefront of science.

BY MONYA BAKER

Elizabeth Iorns did not expect to be attacked for following the scientific method. As a postdoc at the University of Miami in Florida in the late 2000s, she spent a year trying to replicate findings that a particular gene worked as a switch for malignancy. After she concluded that the original research was flawed — a conclusion denied by its authors — she wanted to warn others. But she struggled to publish her data, and faced personal attacks and career setbacks after she succeeded. “It’s still an experience that I don’t like to bring up,” she says.

Given the obstacles, she began to suspect that many published experiments are never repeated. And this year, she decided to do something about it. The need was clearer than ever, she says. Increased retraction rates had drawn the attention of the US National Academy of Sciences. Scientists at Amgen and Bayer reported that they had been unable to reproduce the vast majority of ‘landmark’ papers describing promising approaches to treat disease. The field of psychology was hauled up for widespread replication problems (see *Nature* **485**, 298–300; 2012).

In August, Iorns founded the Reproducibility Initiative, based in Palo Alto, California, which allows authors to submit their papers for replication. Scientific advisers select key experiments and arrange for disinterested third parties to repeat them. If the results are replicated, the validation study is published in *PLoS ONE*, linked to the original paper. This, says Iorns, can draw more attention to the original findings and motivate authors to put their results up for testing.

The Reproducibility Initiative is a limited solution to an important problem, says Ferric Fang, a molecular biologist at the University of Washington in Seattle who has investigated causes of sloppy science. “The ability of work to be reproduced is a hallmark of good science, but it’s difficult to publish work that replicates someone else.” Fang advocates finding easier ways for scientists wishing to expand on the original research to report their replication attempts.

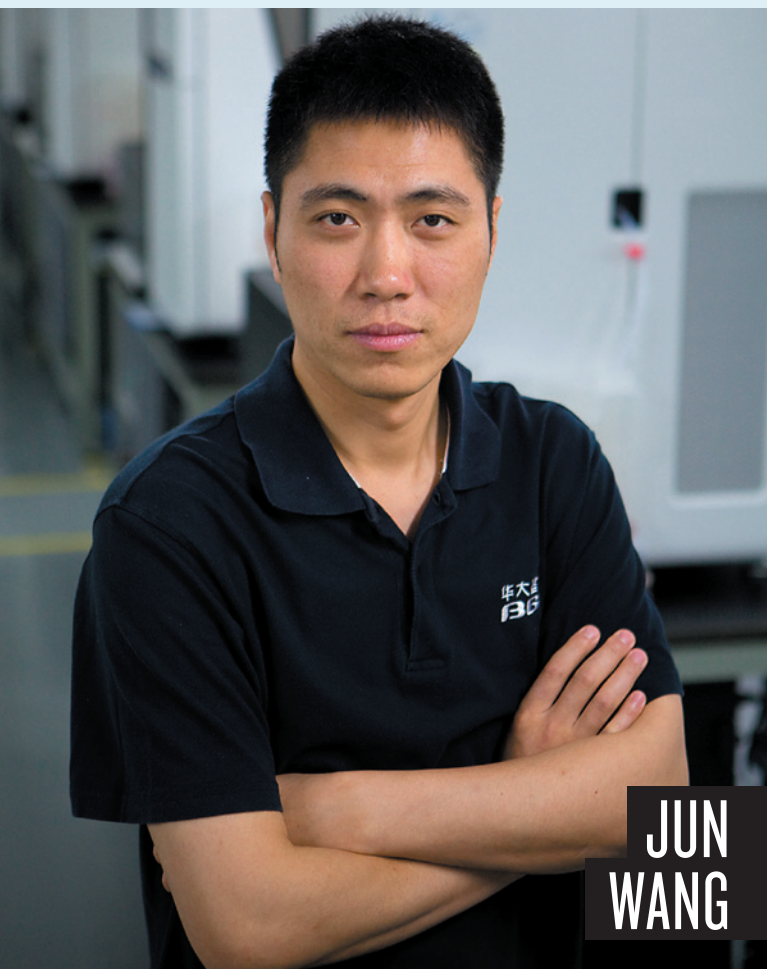
On 14 November, the Reproducibility Initiative e-mailed about 7,000 corresponding authors of recent publications in PubMed asking them if they would like their study validated. By the next day, they had 675 replies, of which 77% said yes. But someone has to foot the bill — which Iorns estimates will be about 10% of the original research cost. Iorns is busy convincing funders that supporting replication work eventually frees up more resources for discovery. “It’s wasteful to not have any idea what is real,” she says. ■

E. J. DYKSON

DELMÍ ALVAREZ/GETTY



CÉDRIC BLANPAIN



JUN WANG

GENOME JUGGERNAUT

The head of a Chinese sequencing powerhouse reveals the scale of the institute's genome ambitions.

BY DAVID CYRANOSKI

“We are the muscle — we have no brain,” said Jun Wang in 2009, describing BGI, the Chinese genome-sequencing institute he leads. It was the kind of cheeky statement for which Wang has become known, but there seemed some truth to the claim. BGI had just purchased 128 top-of-the-line DNA-sequencing machines and was putting hundreds of young programmers — often plucked directly from universities — to work on an onslaught of data. It was a sequencing centre on steroids, using its unparalleled technical power to tackle almost every project going. Today, BGI is the biggest genome-sequencing operation in the world, and Wang, a 34-year-old bioinformatician who has been with it from the start, is well on his way to demolishing the organization’s brawn-without-brains reputation.

BGI was established in 1999 to support the Human Genome Project. It went from accounting for about 1% of the genomics community’s sequencing capacity at that time, to “more like 50%” today, says George Church, a geneticist at Harvard University in Cambridge, Massachusetts. With eight overseas offices and 600 representatives around the globe, BGI works with more than 10,000 collaborators from universities, pharmaceutical and agricultural companies and other research institutions. Although it calls itself a non-profit organization, BGI has courted investors and made moves to acquire other companies in the sequencing field. The rapid growth has been stressful at times. Wang remembers chucking

THE BIAS DETECTIVE

With an experiment that exposed sex discrimination, a microbiologist has opened researchers’ eyes to their unconscious biases.

BY HELEN SHEN

Ever since she saw the results of her study, Jo Handelsman has thought twice about any recommendation letter she writes. Has she somehow slighted a woman or given an unconscious boost to a man, despite her commitment to advocate for women in science?

Last summer, Handelsman injected new life into the long-standing debate about what holds women back in science with a study that showed that both male and female researchers tend to rate job applications from women lower than those from equivalent men (C. A. Moss-Racusin *et al. Proc. Natl Acad. Sci. USA* **109**, 16474–16479; 2012).

Handelsman, a microbiologist at Yale University in New Haven, Connecticut, asked more than 100 scientists to evaluate

applications from undergraduate students seeking a job as a laboratory manager — often a stepping stone to graduate school. Unbeknown to the researchers, the students were fictitious. But the prejudice that Handelsman uncovered was not. On average, researchers who received ‘John’s’ resume said that they would offer an annual salary of US\$30,238; those who read an identical resume from ‘Jennifer’ offered just \$26,508.

Chemists, physicists and biologists — men and women alike — also rated Jennifer as less competent than John and expressed less interest in mentoring her. “There was simply one treatment and one variable, and there’s no escape from the conclusion,” says Handelsman. This type of bias could be one factor holding back female scientists, she says.

Handelsman says that she has never personally experienced significant hurdles because of her sex. But she became an outspoken campaigner for women in science in the 1990s, after hearing how female colleagues and students had faced gender discrimination and seeing them passed up for honours that went to male scientists who were no better qualified.

A wealth of social psychology data had convinced Handelsman that unconscious bias presents a major obstacle for female scientists. But when she spoke about that bias at universities around the country, she encountered scepticism. “I had heard so many times from scientists that this couldn’t possibly be true of us, that we’re trained to be rational,” says Handelsman. So she decided to put it to the test.

“This is one of us — telling us that there are problems,” says Ronald Breaker, who chairs Handelsman’s department. “It comes with a certain amount of street cred.” ■

a computer out of the window during the race to sequence the rice genome a decade ago because his team didn't know enough computer programming for the task at hand. "I had a bad temper. But that's history," Wang says. "I'm very nice and gentle now."

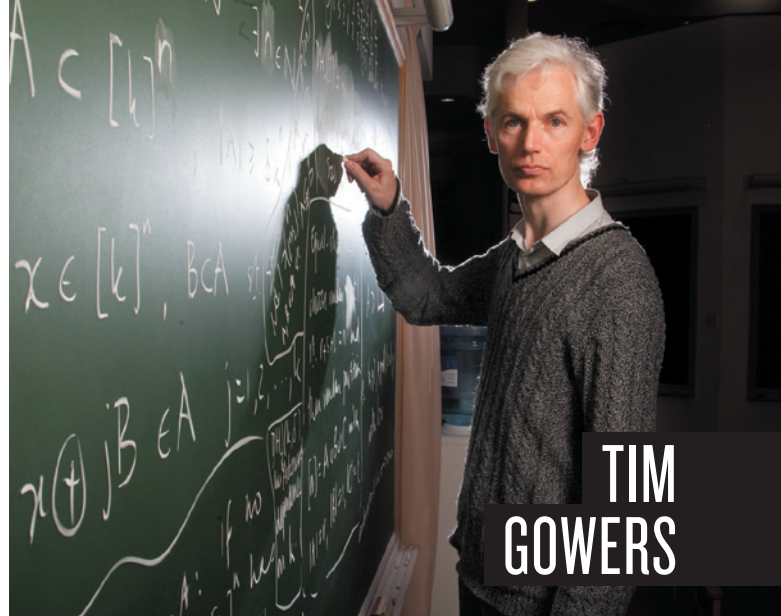
As a public face for the institute, Wang uses his energy and self-effacing humour to highlight BGI's ambitions, which seem to include sequencing the genome of just about every organism on the planet. It is taking a leading role in sequencing 10,000 vertebrates through the Genome 10K project; 5,000 insects and other arthropods through the i5k initiative; and more than 1,000 birds, including some extinct ones in a separate project.

This year, BGI was listed in more than 100 publications. It was a main player in the 1,000 Genomes Project Consortium, which aims to tease out genetic factors in disease by comparing human genomes from geographically distinct regions. And it has increasingly been initiating its own projects, including two studies that analyse the genomes of single cells to chart cancer development (Y. Hou *et al. Cell* **148**, 873–885; 2012 and X. Xu *et al. Cell* **148**, 886–895; 2012).

But the largest change in 2012 was BGI's progress in translating genomic science into real-world applications. A partnership with the Gates Foundation signed in September will expand the repertoire of sequenced agricultural organisms and infectious diseases. The institute is also working on genetic tests that detect fetal chromosome abnormalities from a mother's blood, and it is pushing to use next-generation sequencing for diagnostic tests in newborns.

In an attempt to secure a dominant position in clinical testing, BGI offered US\$118 million in September to acquire sequencing-technology company Complete Genomics of Mountain View, California. Church, who serves as an adviser to both companies, says that Complete Genomics has technologies that will be invaluable in screening for disease-related genes.

Wang says now that he never really doubted that BGI was more than just brawn. "I was just being modest," he says when reminded of his comments from three years ago (see *Nature* **464**, 22–24; 2010). "If you really don't have a brain, you can't move the muscles." ■



**TIM
GOWERS**

SEED OF DISCONTENT

A disgruntled mathematician ends up sparking a global publishing boycott.

BY RICHARD VAN NOORDEN

Tim Gowers is still surprised that he ended up leading a global boycott of Elsevier, the Dutch publishing giant. "I'm not someone who naturally seeks to be a campaigner," says the mathematician at the University of Cambridge, UK. Yet Gowers' impatience with the publisher's business practices had been building for years. He particularly disliked what he saw as its high prices, its habit of forcing libraries to subscribe to unwanted journals by 'bundling' them with the popular ones and its opposition to open-access publishing.

On 21 January, Gowers let rip in a searing blogpost entitled 'Elsevier — my part in its downfall'. "Why do we allow ourselves to be messed about to this extraordinary extent," he wrote. "I am not only going to refuse to have anything to do with Elsevier journals from now on, but I am saying so publicly," he went on — and he encouraged others to do the same.

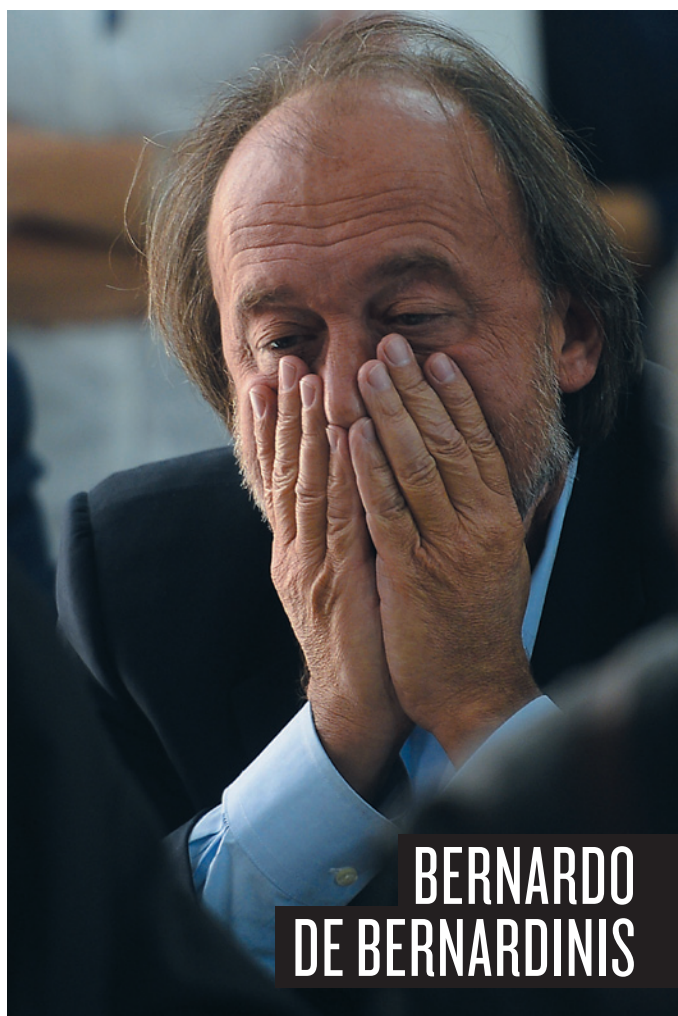
The blog caught the attention of Tyler Neylon, a software engineer in Mountain View, California, who the next day created a website (www.thecostofknowledge.com) inviting people to sign up for a boycott. More than 13,000 scientists across the world have now pledged variously not to publish with or referee or do editorial work for Elsevier.

The signatories are only a tiny fraction of the world's researchers — but the campaign was a spark in this year's explosion of interest in open access and new visions for research publishing. In February, Elsevier, facing growing criticism, withdrew its support of the Research Works Act, a proposal to prohibit the US government from requiring open-access publication for the research it funds. In July, the UK government mandated that much of the nation's taxpayer-funded work be published openly from April 2013. This year also saw the birth of several experimental publishing models, such as the open-access venture PeerJ, which will publish any number of an author's papers for a one-off fee. "We've disagreed with a lot of what [Gowers] has claimed, but for certain he's helped us better understand the sentiment of the maths community," says Tom Reller, head of Elsevier's global corporate relations.

Gowers' campaign was "a little bit accidental," says Ben Green, a fellow Cambridge mathematician. Gowers agrees. He may have won the Fields Medal in mathematics and, this year, a knighthood, but in academic publishing he says he is an amateur. "I feel more like an individual whose views just happened to resonate with others." ■



**JO
HANDELSMAN**



**BERNARDO
DE BERNARDINIS**

ON THE FAULT LINE

Convicted of manslaughter after a deadly earthquake, an Italian official says that he had put his trust in scientists.

BY NICOLA NOSENGO

Minutes after hearing himself declared guilty of manslaughter, Bernardo De Bernardinis looked like a defeated man. With shadowed eyes, he told journalists outside the court in L'Aquila, Italy, that he was "innocent before God and men".

De Bernardinis's journey to court began in early 2009, when a string of seismic shocks around L'Aquila rattled the local community. On 31 March, De Bernardinis and six of the country's seismic experts participated in a meeting of Italy's Major Risk Commission to assess whether a serious earthquake might be imminent. Some townspeople say that the advice they received from the experts was reassuring — but six days later, a magnitude-6.3 quake struck, killing more than 300 people in the region. In October this year, the court of L'Aquila found the seven guilty of the manslaughter of 29 of those people, and sentenced them all to six years in prison.

The trial and verdict attracted worldwide attention — and De Bernardinis was at the centre of the storm. An engineer by training, De Bernardinis was in 2009 deputy head of the Italian Department of Civil Protection. He was the only government official on the expert panel and was charged with deciding what to do. In a now-infamous television interview shortly before the meeting, he said that the situation was "favourable... according to the scientific community", that minor shocks were linked to "a continuous discharge of energy" and that "there is not an immediate danger".

The trial prosecutors argued that those messages falsely reassured the local population. Seismologists from the National Institute for

FLIPPON MONTEFORTE/AF/GETTY

FLU FIGHTER

A virologist did research that some deemed too dangerous to publish, and he spent much of the year defending his work.

BY DECLAN BUTLER

The Dutch are known for being blunt, and Ron Fouchier, a virologist at the Erasmus Medical Centre in Rotterdam, certainly seems to live up to that reputation. His candour caught global attention at the end of 2011, when he described his experiments to engineer a strain of highly pathogenic H5N1 avian influenza that can be transmitted between mammals. His team had "mutated the hell out of H5N1", Fouchier told a flu-research conference in Malta, until it could infect ferrets — and presumably humans — through the air. Just five mutations were enough to cause this change, suggesting that wild H5N1

strains might eventually evolve to spread in humans. "This is very bad news, indeed," he said.

His work (S. Herfst *et al. Science* **336**, 1534–1541; 2012), and similar experiments by Yoshihiro Kawaoka (M. Imai *et al. Nature* **486**, 420–428; 2012), a virologist at the University of Wisconsin–Madison and the University of Tokyo, prompted the US National Science Advisory Board for Biosecurity (NSABB) to recommend in December 2011 that the research findings be published only if key methodological details were left out. The events sparked an international debate as to whether the risks of an accidental or intentional release of the pathogens outweighed any benefits of the research.

Fouchier relentlessly defended the rationale and safety of his work. In public debates, commentaries and media interviews, he often dismissed opponents' arguments against the research, playing down the risks, and he often seemed exasperated when arguing against those who scrutinized the flu-research community. In March, the NSABB reconvened and eventually gave the green light to publishing both papers. Fouchier was at the airport on his way home from the NSABB meeting when he heard the news. "Boy, did that glass of champagne in the airplane taste good!" he recalls.

Still, the Dutch government refused to let him publish until he applied for an export permit, which the government requires for the

Five to watch

2013

ANNE GLOVER

EUROPEAN COMMISSION CHIEF SCIENCE ADVISER

As the first person to hold the position, Glover is planning to tackle hot-button topics such as genetically modified organisms and the use — or abuse — of science in policy-making. Let the fights begin.

THOMAS STOCKER

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE

As co-chairman of the IPCC's Working Group I, which will publish its report on the physical science aspects of climate change in 2013, Stocker will channel a familiar but increasingly urgent message.

CHRIS AUSTIN

US NATIONAL CENTER FOR ADVANCING TRANSLATIONAL SCIENCES

There are big claims, high expectations and sharp critics; now the head of the US National Institutes of Health's bold new venture has to show that academia can succeed at drug discovery where pharma has struggled.

JAN TAUBER

EUROPEAN SPACE AGENCY'S PLANCK MISSION

When Tauber's team releases the most precise map yet of the cosmic microwave background, astronomers will comb the data for evidence of gravitational waves associated with inflation: a wrenching expansion of the Universe hypothesized to have occurred just after the Big Bang.

RAFAEL YUSTE

COLUMBIA UNIVERSITY, NEW YORK

Big neuroscience is in vogue and nothing comes bigger than Yuste and his colleagues' proposed Brain Activity Map Project, which aims to record all electrical activity from every neuron in a circuit.

Geophysics and Volcanology (INGV) in Rome, some of whom were tried alongside De Bernardinis, have criticized him for suggesting that minor shocks release energy and lower the risk of a big earthquake — an incorrect concept, they say, which was not raised at the meeting or in any official INGV communication.

Throughout the trial, De Bernardinis, now president of the Institute for Environmental Research and Protection in Rome, never hid from his critics. He was the only one of the indicted who showed up at every hearing and he speaks with respect of the citizens who brought the case. "Had I lost a son, a relative or a friend in the earthquake, I would have done the same," he says.

In November, a few weeks after the verdict, De Bernardinis is ready to resume battle at appeal. A chatty, energetic man in his mid-60s, he flips through hundreds of pages of press clippings and documents that he has put together for the case. He pounds his fist on the table when he rejects charges that he meant to reassure the population or that he denied the risk of a big earthquake. In interviews before and after the L'Aquila meeting, he insists, he was "repeating, not modifying" concepts used by the seismologists, and he denies saying that the discharge of energy

decreased risk. He blames the press for misreporting his interviews and says that by "favourable", he meant that the minor shocks had not yet caused major damage. De Bernardinis does acknowledge that he should have waited until after the meeting before giving an interview, and that he would have avoided problems if he had asked the scientists to prepare a written note for the press.

"I understand scientific language, but I am not a seismologist," he says. "I could only trust what seismologists said." The prosecutor at the trial, Fabio Picuti, seemed to agree when he called De Bernardinis "a victim of the seismologists" in his final argument; the seismologists disagree.

While the case inches towards appeal, De Bernardinis continues to work, and keeps the wrinkled trial documents in his briefcase. "They follow me everywhere," he says.

De Bernardinis hopes that the trial will eventually lead to a better risk-prevention system in Italy, by clarifying the obligations of scientists, government officials and the media. He still considers himself innocent. "But if at the end of the appeals I will still be found guilty, I'll go to jail, no problem," he says. "I'd rather go to jail feeling I am innocent than stay out feeling I'm guilty." ■

LEVEN WILLEMS

release of material or information that could have malicious uses. Fouchier threatened to defy the mandate, but backed down after it became clear that such action risked up to six years' imprisonment (see *Nature* <http://doi.org/hvg; 2012>). He got the permit and published the work, but is still disputing in the courts that export-control laws apply. "If there is anything I can do to prevent future generations of infectious-disease specialists being censored against their will by government, I will do it," says Fouchier.

In January 2012, flu biologists including Fouchier agreed to a voluntary moratorium on the specific kinds of experiments that had sparked the debate. Although initially set for 60 days, the moratorium remained in place as *Nature* went to press, and Fouchier, among others, complains that it has gone on for too long. It has put a hold on some of his research, but Fouchier says that a mysterious case of deadly pneumonia in Saudi Arabia over the summer drew him back into the lab. He discovered that the victim had died from a previously unknown coronavirus that has since been recognized as the cause of an outbreak in the Middle East (see *Nature* 492, 166–167; 2012). Before that case, Fouchier had been resigned to spending the year dealing with the "politics, censorship, moratorium, biosafety and biosecurity" of the H5N1 experiments, he says. "It was a good therapy for me to get back to work." ■



COMMENT

PUBLISHING Open access must enable reuse to drive innovation **p.348**



VITICULTURE Making wine in a warming world **p.351**

EXHIBITION Treasures and their tales highlight Natural History Museum's trove **p.354**

HEALTH Rousing polemic details the dangers of too much sugar **p.355**

CORE/IS



J. Robert Oppenheimer (third from left) and General Leslie Groves (centre) examine the remains of the first successful atomic bomb test in 1945.

Self-censorship is not enough

The debate over publishing potentially dangerous research on flu viruses would benefit from a closer look at history, argue **David Kaiser** and **Jonathan D. Moreno**.

After being mutated just a handful of times, an artificially created variant of the H5N1 avian influenza virus began to spread among ferrets in virologist Ron Fouchier's laboratory in Rotterdam, the Netherlands, late last year. This was ominous, because ferrets are a model for human-to-human transmission of flu.

Fouchier's work, along with parallel studies by virologist Yoshihiro Kawaoka of the

University of Wisconsin–Madison and the University of Tokyo, ignited a worldwide debate. Earlier this year, experts in microbiology and in science policy wrestled over whether the specifics of such potentially dangerous research should be openly published in scientific journals¹.

Some insisted that publishing detailed information on how to produce mutant flu strains is crucial to researchers who want to

fine-tune surveillance strategies for pandemic outbreaks and hone potential vaccines. Others argued that releasing this information will increase the chance of a deadly virus devastating human populations — either by enabling bioterrorists to manipulate strains, or by aiding an accidental release from a lab. Many even questioned whether such research should have been funded in the first place.

The furore, which is far from settled ►

► (see go.nature.com/wxeijg), has sparked passionate claims and counterclaims about the risks of imposing tight controls over the flow of biological information. Missing from the discussions has been a clear-eyed look at history.

The problems posed by dual-use research — which can benefit the public but might also be co-opted for harmful purposes — are hardly new. Even in ancient Greece, Archimedes applied his mathematics to improve devices used to overcome city walls in siege warfare. Revisiting the successes and mistakes of the recent past would clarify the risks and benefits of various proposals for biomedical research. We argue that, although scientists are right to be wary of heavy-handed approaches to security, self-censorship alone has rarely proved sufficient.

FLU FREEZE

Reacting to the wrangling over whether the H5N1 work by Fouchier and Kawaoka should be published, virologists adopted a voluntary moratorium on flu-virus research in the early months of 2012. Several practitioners have argued that such self-governance is adequate for today's challenges in the life sciences, and that further checks on the open sharing of information would hamper scientific progress^{1,2}. This is surprising to us: no matter the field of research, can anyone be expected to step outside the excitement and momentum of their own work to make objective decisions in risky situations?

Biologists' resistance to meddling from outsiders probably stems, at least in part, from fears of the kinds of restriction and paranoia that constrained many nuclear physicists working in Europe and the United States during and after the Second World War.

After the onset of the war, US nuclear physicists agreed among themselves in the early 1940s to restrict or avoid publication of certain information about nuclear reactions, even before the Manhattan Project³ to develop the atomic bomb was established. But formal schemes for classifying such information as secret were soon codified and written into law. The US Atomic Energy Act of 1946, for instance, made it a federal crime to circulate information about the rates of some nuclear reactions without extensive review and declassification. This ensured that whole categories of information were deemed secret. The default became to classify first and to declassify only when needed.

Many physicists at the time chafed at the new rules. The restrictions were also ripe for abuse. Historians have uncovered dozens of instances in which authorities in the United States, the United Kingdom and elsewhere used secrecy and classification to undermine legitimate discussion and debate. In the United States, for example, more physicists were called to testify before the House

Un-American Activities Committee than members of any other academic discipline⁴.

Classification of information was sometimes used unfairly to disadvantage defendants. This happened in personnel security hearings (including, most famously, that of J. Robert Oppenheimer, scientific director of the wartime Los Alamos Laboratory in New Mexico), in criminal espionage cases (such as that of accused atomic spies Julius and Ethel Rosenberg) and even in intellectual-property claims. Nor was the system foolproof. There

"Asilomar has been cited as a shining example of scientists' ability to act responsibly."

were several instances of espionage at Los Alamos despite the elaborate security precautions⁴.

This monolithic regime of classification, emblematic of the Second World

War and the post-war 'red scare' era of anti-communism, however, is one of many examples of concerned non-scientists stepping into the research fold.

RECOMBINANT RESEARCH

In 1975, molecular biologists met at the Asilomar Conference Grounds in Pacific Grove, California, to discuss scientists' concerns about the emerging field of recombinant-DNA technology. A major worry at the time was that introducing genes from, say, viruses could transform innocuous microbes into deadly pathogens that were resistant to antibiotics or into agents that cause cancer. The biologists agreed to a voluntary moratorium on all research involving recombinant DNA until they could hammer out appropriate safety protocols⁵.

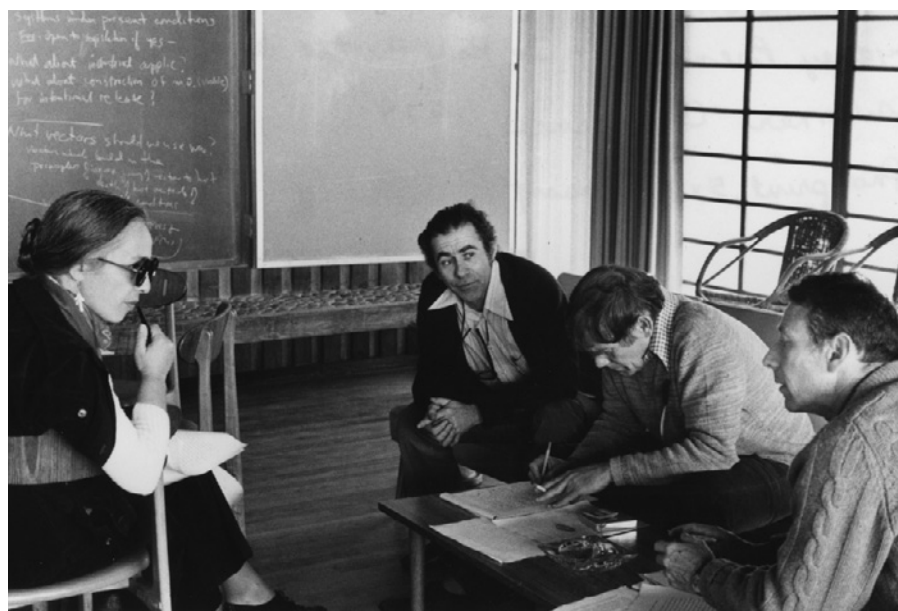
In the recent debate over work on flu viruses, Asilomar has been cited as a shining

example of scientists' ability to act responsibly when unfettered. Yet self-censorship was only the beginning of that chapter.

Around the time of the Asilomar meeting, local officials in US cities — notably Cambridge in Massachusetts, but also Ann Arbor in Michigan, Bloomington in Indiana, Berkeley and San Diego in California and Madison in Wisconsin — also became concerned about the potential risks of recombinant-DNA research. Rather than rely on the scientists to police themselves or wait for the US National Institutes of Health to dispatch safety guidelines, the mayor of Cambridge, Alfred Vellucci, convened public hearings on the topic. He even threatened to impose a ban on researchers to stop them from carrying out any recombinant-DNA research within city limits unless they first worked with multiple local stakeholders to iron out mutually agreeable safeguards⁶.

Some molecular biologists at the local powerhouse institutions — Harvard University and the Massachusetts Institute of Technology — grew frustrated by the imposition. Following four weeks of debate, and after the city had imposed a three-month moratorium on all recombinant-DNA research, the Cambridge Experimentation Review Board was established in 1976 to consider the public-health effects of the research. The board's members included physicians, academics from other disciplines and members of the public. They met twice weekly for five months, arranged public debates and considered testimony from dozens of experts and non-specialists — including Nobel laureates, university and public officials, and concerned citizens.

Ultimately, the board crafted city legislation establishing a level of cooperation among researchers, policy-makers and



Maxine Singer, Norton Zinder, Sydney Brenner and Paul Berg (left to right) at the 1975 Asilomar meeting.

US NATL LIB. MED.



Nuclear research, such as that at Oak Ridge National Laboratory, is harder to hide than bench-top biology.

concerned citizens that was unprecedented for a local effort. The bill — passed unanimously by the city council in 1977 — provided a model for other cities, and similar local negotiations unfolded elsewhere at around the same time⁶.

With regulatory uncertainties removed, Cambridge especially became a magnet for investors who sparked the new biotechnology industry off the back of recombinant-DNA research. Companies such as Biogen (now Biogen Idec), Genzyme and Millennium Pharmaceuticals were launched⁶.

DUAL-USE DILEMMA

What aspects of these two critical moments in history — the Asilomar meeting and the handling of post-war nuclear physics — are relevant to dual-use research in the life sciences today?

Arguably, despite the downsides, a certain clarity emerged from nuclear scientists' rules of conduct being written into national laws during the 1940s and 1950s. Grey areas over what was permissible and who had ultimate jurisdiction were kept to a minimum. Yet it is difficult to see how an approach modelled on the US Atomic Energy Commission, which oversaw the peacetime development of nuclear technology until its dissolution in 1975, would fit today's challenges for the life sciences.

Most of the sensitive research related to nuclear technologies has required hulking, factory-sized infrastructure that shows up in grainy satellite images from space. By contrast, much of the biological research that could have dual uses today — from studies of drugs that alter memories to genetic manipulations of viruses or bacteria — is conducted in groups of just a few people, requiring bench-top equipment.

There are also important differences

in patronage and jurisdiction. The huge facilities required for nuclear projects, such as the Oak Ridge National Laboratory in Tennessee, mean that most high-risk work has been conducted under the auspices of a national government. Such research has been underwritten almost exclusively by federal funds, and thus is naturally subject to government oversight and control.

Today, many biologists are funded by multinational corporations or philanthropic organizations in addition to federal grants. This has blurred matters of jurisdiction, oversight and ownership of scientific results⁷. In 2005, for example, the *Proceedings of the National Academy of Sciences* published an article⁸ that analysed the possible effects of botulinum toxin contamination of the US milk supply. The journal's decision to publish was explicitly opposed by the US Department of Health and Human Services⁹.

"Few physical scientists in the 1940s and 1950s argued that self-censorship would suffice."

In their paper, the researchers acknowledged funding from Stanford University's business school but none from federal sources. The journal's editors concluded that federal overseers did not have sufficient jurisdiction to block the article's publication.

Some policy statements, such as a 2007 report from the US National Research Council¹⁰, have harked back to the Asilomar meeting and called for renewed self-regulation in the life sciences. Such reports acknowledge that internal institutional review boards have often failed to assess dual-use bioscience adequately in recent years, and have pushed for more education of life scientists "on the basic principles of risk-based biosafety and biosecurity review"¹⁰. Recognizing the need

for some government input and leadership, some reports have suggested that bodies such as the US National Science Advisory Board for Biosecurity (NSABB) provide advice.

In the 1970s, however, workable systems of oversight and security emerged from intensive engagement with a broad range of legitimate stakeholders. The recent H5N1 debacle in particular raises questions about whether self-regulation with guidance from a government body is all that different from self-censorship alone.

Initially, the NSABB voted for the US federal government to block publication of portions of Fouchier's and Kawaoka's papers. A few months later, the group reversed its decision, supporting full publication. Several experts testified before the US Senate that they feared the NSABB had stacked the deck by soliciting input predominantly from the very scientists who had conducted the work in question (see go.nature.com/wocmeh).

Amid the drama and passions roused by the H5N1 saga this year, it is worth remembering that few physical scientists in the 1940s and 1950s argued that self-censorship would suffice — despite the seemingly draconian approach of the Atomic Energy Commission. Indeed, when asked why they had continued to work on the atomic bomb after Germany had been defeated, most of the physicists at Los Alamos admitted to having been caught up in the momentum of the project. In the 1970s, biologists were not relied on to regulate matters of much broader concern on their own, despite their much-heralded initial efforts to self-govern. Moreover, the involvement of many other players did not quash scientific research and communication, but helped molecular biology and biotechnology to thrive. ■

David Kaiser is professor and department head of the Program in Science, Technology, and Society at the Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. **Jonathan Moreno** is professor of medical ethics, history and sociology of science, and philosophy at the University of Pennsylvania, Philadelphia, USA. e-mails: dikaiser@mit.edu; morenojd@upenn.edu

1. Maher, B. *Nature* **485**, 431–434 (2012).
2. Butler, D. *Nature* **486**, 449–450 (2012).
3. Weart, S. R. *Phys. Today* **29**, 23–30 (1976).
4. Kaiser, D. *Representations* **90**, 28–60 (2005).
5. Berg, P., Baltimore, D., Brenner, S., Roblin, R. O. & Singer, M. F. *Proc. Natl Acad. Sci. USA* **72**, 1981–1984 (1975).
6. Durant, J. in *Becoming MIT: Moments of Decision* (ed. Kaiser, D.) 145–163 (MIT Press, 2010).
7. Mirowski, P. *Science-Mart: Privatizing American Science* (Harvard Univ. Press, 2011).
8. Wein, L. M. & Liu, Y. *Proc. Natl Acad. Sci. USA* **102**, 9984–9989 (2005).
9. Kaiser, J. *Science* **309**, 31 (2005).
10. National Research Council Science and Security in a Post 9/11 World (National Academies Press, 2007).

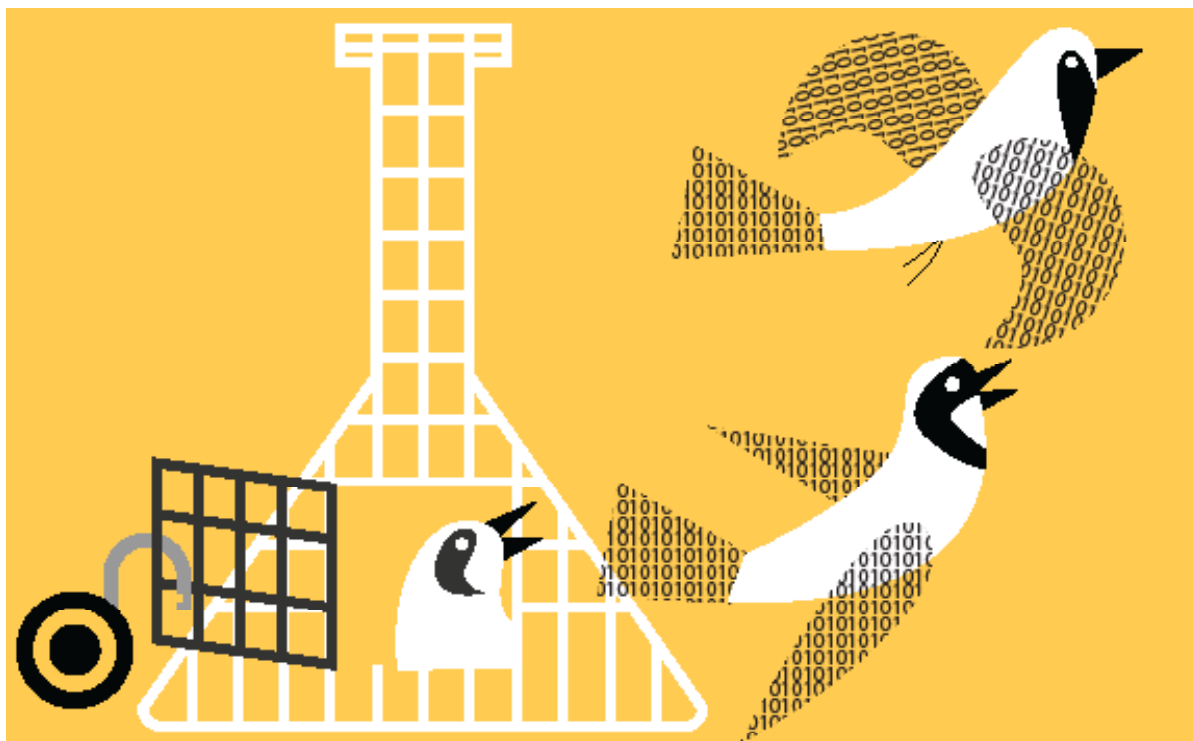


ILLUSTRATION BY MATTHEW HOLLISTER

Open access must enable open use

Those wishing to maximize the benefits of public research must require more than free access, says **Cameron Neylon** — they must facilitate reuse.

At the end of 2012, there is no longer a debate about whether the results of research should be freely accessible. All that remains is to work out when and how access will be enabled. Meanwhile, the political and economic question of 2012 has been: should governments invest to nurture economic recovery, or tighten their belts and risk further economic damage? Publicly funded research has often been at the heart of this debate as governments attempt to ensure that public investment is generating the greatest possible innovation, economic activity and societal gains.

It is in this light that researchers should view the dramatic advance of open access in 2012. This shift, and the reason why governments and funders are increasingly adopting or examining open-access mandates, is about more than just the benefits of access. It is about the potential of open access to improve the efficiency of research itself, and to deliver a greater return on public investment.

In this context, free access is not enough. To maximize the value of the public good

and the return on investment, research outputs must be reusable. That does not just mean making data or papers available on the Internet, but ensuring that innovators can manipulate the material, including mining, translating or expressing it in imaginative ways or for new audiences.

The most significant policy moves in open access in 2012 came from Research Councils UK (RCUK) and the Wellcome Trust, a major UK biomedical funder (see go.nature.com/qqqrsq and go.nature.com/u2oqym). From April 2013, papers arising from RCUK funding must be made free to access. Significantly, the core of the policy aims to maximize the reusability of research outputs, not just provide free access. Building on its long-standing leadership in this area, the Wellcome Trust strengthened its open-access policy, harmonizing it with that of the RCUK.

Making research outputs usable has many aspects. The technical side — standardizing the representations of data and knowledge in ways that make them easily transferable — remains challenging and needs further work.

There are also legal challenges, but good tools exist that provide the rights to reuse research in any way that scientists can imagine: the Creative Commons licences. These easily and effectively define precise 'reuse rights'.

The Creative Commons Attribution (CC BY) licence — a key component of the RCUK and Wellcome policies on open access — allows any kind of reuse, provided the copyright holder is properly attributed. PLOS, where I work as advocacy director, has always used the CC BY licence, like other major open-access publishers, and this licence is emerging as a standard for open-access publications. The RCUK and Wellcome policies explicitly encourage researchers to submit their papers to journals that will publish the work under a CC BY licence — making these the first funders to take such a step. The target is to create a critical mass of freely accessible and freely reusable literature. The RCUK and Wellcome have taken the lead, now others must follow if everyone is to reap the full benefits of effective research communication.

There are two established mechanisms

for making research articles accessible. One is for the publisher to make the version of record freely available online. The second is for a version to be made available in an institutional, disciplinary or funder repository. These two routes are often referred to as 'gold' and 'green' (see go.nature.com/ao7k4b). However, because 'gold' has been misused in the past, I will refer to journal-mediated and repository-mediated open access, rather than gold and green, respectively.

ACCESS FOR ALL

Repository-mediated open access has provided the greatest volume of accessible material in the past ten years. Some analyses have estimated that more than 20% of the world's peer-reviewed literature can be read for no charge through this mechanism¹. One of the major contributors to this growth has been the Public Access Policy of the US National Institutes of Health (NIH), which requires that either the published version or the final author copy, after peer review, of NIH-funded papers be deposited into the online archive PubMed Central within 12 months of publication. The policy, along with others, has provided access to an archive that so far contains around 2.5 million articles.

The success of PubMed Central and of other disciplinary and institutional repositories illustrates a weakness. Although millions of articles are accessible to read, the majority of them cannot be used for anything except reading. If, for instance, you wish to index all the gene names in a set of papers, put them on a website, translate them, use text or images in a summary or even just print out several copies of the collected papers, you are limited to a much smaller set of around 500,000 articles that carry a Creative Commons licence (see go.nature.com/heaqoe). For any commercial purpose, which could include simply making copies for a class or company meeting, one is restricted to the smaller subset of papers that have a CC BY licence.

This matters because a growing body of evidence shows that enabling the reuse of research maximizes its potential for innovation. A natural experiment conducted in the academic community, in which an agreement with the NIH made one set of mouse strains for genetics research more freely accessible, showed that these strains generated more citations, as well as more diverse and more applied research². The publicly funded Human Genome Project and its freely reusable data generated a massive 141-fold return on investment in economic returns alone³. In terms of its wider impact, the reusable data led to 30% more new clinical products than the privately funded, closed genome-sequencing project of the US biotech firm Celera Genomics⁴.

There are fewer and smaller examples that relate to the published research literature, because a critical mass of accessible and

reusable papers has not yet been achieved. This is what the RCUK and Wellcome policies seek to achieve.

Unfortunately, funders cannot solve issues of reusability simply by requiring papers to have a CC BY licence. When funders are not covering publishing fees, it is more difficult for them to require specific licensing terms. The RCUK and Wellcome both allow researchers to use the repository route to open access but, although they request certain terms of reuse, they do not demand it. Traditionally, authors hand control of reuse rights to publishers. For this to change, funders need to gain leverage with publishers.

One mechanism is to buy that leverage. When a funder is paying for publication services, they are in a position to dictate service levels — enabling the RCUK and Wellcome, for example, to require that papers for which they fund publication charges carry a CC BY licence. A consortium approach to creating a journal-mediated open-access environment for particle physics, SCOAP3 (<http://scoap3.org>), achieved the same ends earlier this year by running a competitive tender process for the publication services they required and making licensing terms part of the requirements.

An alternative approach builds on the NIH policy, which stipulates that funded authors retain rights to their research so that it can be placed in a repository. There is no legal reason why this could not be expanded to depositing a copy with a specific licence. This would make it possible to use repository-mediated open access to provide reusable content with the right legal tools.

THE POLITICS OF PROGRESS

The decision in the United Kingdom has been that the benefits of access and reuse will be achieved more quickly and efficiently by supporting the growth of the journal-mediated route. Fundamentally, this is a political judgement. Funders do not feel that they have the political leverage to take a rights-retention approach. The RCUK and Wellcome have decided that the most effective way to support progress towards open access and reuse is to provide the resources to support journal-mediated access.

Not all agree that the focus on licensing is the correct priority. Proponents of repository-mediated open access point to the greater pool of research currently available through repositories, and argue that the cost of delivering that access is lower. In my view, funders who want to maximize the return on their investment in research communication should use the leverage provided by

"Institutions need to negotiate more imaginative and favourable arrangements with subscription publishers."

financing to require immediate access to the final published version, with the right licensing. The RCUK and Wellcome policies are the first to actively do this and represent a significant step forward.

Much of the criticism of the RCUK policy has focused on the transitional costs — while a mix of open access and traditional publishing persists, it will be necessary for institutions to pay both subscriptions and publication charges. However, simple 'back of the envelope' calculations based on current charges by open-access publishers⁵ and an estimate of a few million papers per year suggest that there are significant savings to be made following a move to full open access. Studies of the situation in the United Kingdom show that all institutions would see savings from this approach⁶ — if so, a few years of investment is just smart management.

More broadly, institutions need to take the opportunity to negotiate more imaginative and favourable arrangements with subscription publishers, to constrain transitional costs. Twenty-one years of near 100% open access through the arXiv for particle physics, for example, has not reduced subscriptions to particle-physics journals⁷. But when the community used its combined purchasing power through the SCOAP3 consortium, an entire field was converted to open access with no additional transitional costs (see go.nature.com/aw4jtq).

No change comes without risks. This year's UK policies are among the first to focus on ensuring that research communication is configured to maximize the benefits of the Internet. Other funders equally focused on encouraging innovation from public research must follow this lead. Licensing is only one part of that optimization, but it is an important element that we can tackle today. ■

Cameron Neylon is advocacy director at PLOS. He writes regularly on the technical and social issues of using the Internet effectively in scholarly communication, and is based in the United Kingdom.
e-mail: cn@cameronneylon.net

1. Swan, A. *Policy Guidelines for the Development and Promotion of Open Access* (UNESCO, 2012).
2. Murray, F., Aghion, P., Dewatripont, M., Kolev, J. & Stern, S. *Of Mice and Academics: Examining the Effect of Openness on Innovation* NBER Working Paper No. 14819 (National Bureau of Economic Research, 2009).
3. Battelle Technology Partnership Practice *Economic Impact of the Human Genome Project* (Battelle Memorial Institute, 2011).
4. Williams, H. L. *Intellectual Property Rights and Innovation: Evidence from the Human Genome* NBER Working Paper No. 16213 (National Bureau of Economic Research, 2010).
5. Solomon, D. J. & Björk, B.-C. *J. Am. Soc. Inf. Sci. Tech.* **63**, 1485–1495 (2012).
6. Swan, A. & Houghton, J. *Going for Gold? Report to the UK Open Access Implementation Group* (UK OAIG, 2012).
7. Swan, A. *Open Access Self-archiving: An Introduction* (Key Perspectives, 2005).

PAUL PICKFORD/ALAMY



VITICULTURE

Fruity with a hint of drought

Jamie Goode tracks how our changing climate is sending ripples of disruption through the wine world.

For anyone who enjoys a glass of Sancerre, Puligny-Montrachet or indeed a bargain-basement European blend, 2012 is unlikely to be a 'good year'. Growers will remember it as one of the worst in recent history. The grapevine, *Vitis vinifera* L., is remarkable among crop plants in its sensitivity to deviations in climate — so much so that wine is proving to be a canary in the coalmine for climate change. Even small shifts in growing-season temperatures show up as marked differences in flavour or yields. And this year, unusual weather patterns in Europe — which accounts for more than 60% of global wine production — have played havoc with that vulnerability.

An extremely dry summer in southerly regions and exceptionally cold conditions in northern grape-growing areas battered yields disastrously. France suffered a 20% drop in wine production. Reports suggest that harvests in the Loire Valley, home of Pouilly Fumé, for example, were down by as much as 50%; and that Champagne and Burgundy saw declines of 40% and 30%, respectively. Meanwhile, estimates for Italy indicate a 7% drop on what was already, in 2011, a reduced crop. Some 2.5 million families in Europe depend on wine for their living, and this year could have dealt the killer blow to many small family-owned vineyards.

Keeping a weather eye on climate is now part of viticulture. In 2005, the wine world was shaken by a study by climatologist Gregory Jones at Southern Oregon University in Ashland and his colleagues. They analysed 50 years of average growing-season-temperature data from 27 important wine regions and compared them with expert assessments of vintage quality (G. V. Jones *et al.* *Clim. Change* 73, 319–343; 2005). They also ran the Hadley Centre climate model to look at the projected temperature changes in key wine regions up until 2049. The trends that Jones' team identified suggest that much of the wine world is in for an uncomfortable ride over the coming decades.

A look at what weather such as drought, drenchings and gales can do to grapes reveals why. If the weather is unsettled during flowering (mid-June in the Northern Hemisphere), yields can plummet because of poor fruit set leading to fewer berries and smaller bunches. By contrast, warm, damp conditions during the growing season can encourage the fungal diseases to which *V. vinifera* is highly susceptible.

Developing grape berries are sensitive to drastic changes in the weather. Too little light and warmth, and they struggle to reach 'sugar ripeness' (with potential alcohol levels of 12–14%), the point at which sugars rise and levels of organic acids decrease. Too much light and warmth, and they reach sugar ripeness before hitting ▶

► physiological, or ‘phenolic,’ ripeness. This means that polyphenolic compounds such as tannins and anthocyanins in the fruit’s skin never undergo natural modifications to smooth out their astringency. Because grapes should be harvested only at phenolic ripeness, growers are then forced to correct problems in the winery — for low acidities, for instance, they might add tartaric acid — which can lead to lower-quality wines.

Wine is, in any case, complex. Jancis Robinson and co-authors’ authoritative *Wine Grapes* (Allen Lane, 2012) describes 1,368 commercially used grape varieties, each with its own characteristics. These range from Cabernet Sauvignon, which can make big, strong, tannic red wines, to Riesling, known for producing delicate, floral, expressive white wines. Determining the vineyards in which each variety will perform well hinges largely on climatic factors and soils: changes in subsoil composition over small distances can have a huge impact on the characteristics of the final wine. The myriad possible interactions of grape variety, climate, growing season and soils means that the array of wines on the market is bewildering — CellarTracker,

a leading database of enthusiast-generated wine-tasting notes, lists more than 1.4 million different wines.

The expansion of viticulture into previously unplanted regions has taught producers much about this interplay of factors. Growers in new regions have had to learn about the influence of different climates on the way grapevines grow, much as they are now having to adapt to climate change. Vineyard prospecting has occurred mainly in the Southern Hemisphere — Australia, New Zealand, Chile and South Africa. Here, intrepid producers have sought out cooler climates, in a bid to emulate the classic fine wines of Europe. For example, some are keenly seeking places where Pinot Noir, the delicate red grape behind elegant Burgundies, might thrive.

Lupin breeder John Gladstones, for instance, sparked a revolution in Western Australia — now home to some of the

“Wine is proving to be a canary in the coal mine for climate change.”

country’s best Cabernet Sauvignon-based reds — by trawling weather and geological data in the 1960s. Gladstones identified the Margaret River area three hours’ drive from Perth as having a Bordeaux-like climate and soils for quality wine growing. Four pioneer wineries duly began to plant vines. And in Chile, new cool-climate vineyards such as those of the Leyda, Elqui and Limari valleys, most of which were established during the past 10 years, are already making interesting wines.

Today’s ideal growing conditions are at the mercy of tomorrow’s climate change, however. Records in Burgundy stretching back to the medieval era show that warm periods similar to that of the 1990s occurred in the 1380s, 1420s, 1520s and 1630s, each time followed by cooler periods. But as Jones and his team in Oregon found, the second half of the twentieth century witnessed a unique trend, with an average rise of 1.26 °C in growing-season temperatures across the world’s wine regions. And there has been no sign of cooling since, with 2003 the hottest year since these records began.

For some places, that trend has been a

Vintage scientists

Nature uncorks five tales of researchers who plunged into the heady science of viticulture.

Physicist Richard Feynman said that “if we look at a glass of wine closely enough we see the entire Universe”. It is the stuff of physics, chemistry, geology, psychology and the ferment of life itself. Indeed, viticulture, oenology and wine have long fascinated scientists, from Alexander Fleming and Louis Pasteur to Galileo Galilei, who described wine as “sunlight, held together by water”. Vineyards and wineries are, after all, laboratories in which the fruit of the grapevine *Vitis vinifera* is biochemically transformed, by sunlight, rain, soils, precision, judgement, time and technology, into a potentially sublime elixir. So, here we raise a glass to some of the adventurous scientists — including a Nobel-prizewinning astrophysicist, a group of engineers and a neuroscientist — who, with exactitude and flair, have made waves in the wine world.

Maipenrai, Canberra District, Australia

One of the starriest viticulturalists has to be astrophysicist Brian Schmidt, who runs the 1.1-hectare Maipenrai vineyard near Canberra. Schmidt’s work, observing type Ia supernovae at distances of more than 5 billion light years (1,533 megaparsecs), helped to prove the accelerating expansion of the Universe — and led to a share in the 2011 Nobel Prize in Physics. His plot, planted 12 years before that win with six Pinot Noir clones, now produces some 3,000 bottles of Maipenrai wine a year. Schmidt’s sold-out 2009 vintage had, in his own words, “lovely aromas of plums mixed with forest floor”, a “rich, slightly savoury palate” and — not unlike the astrophysicist’s ever-expanding findings on the day job — “a long life ahead of it, gaining complexity”.



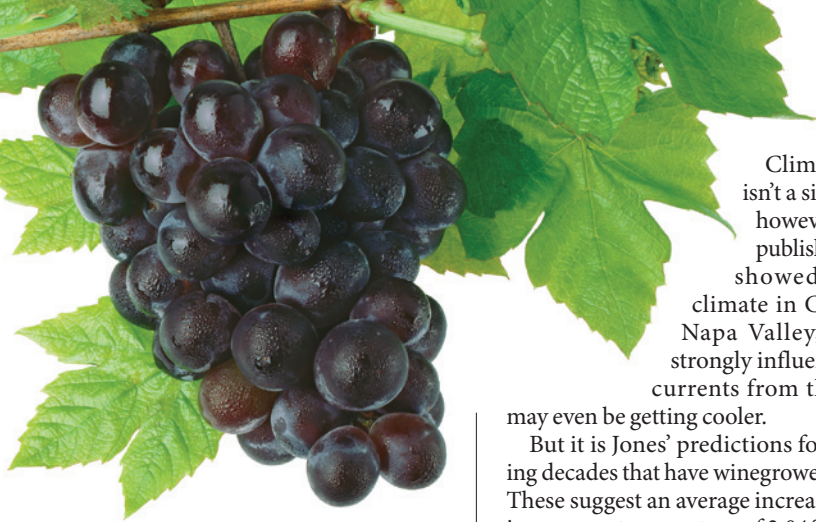
Lethbridge, Geelong, Australia

For 10 years, neuroscientist Ray Nadeson worked on the pain-relieving properties of neurosteroids at Monash University in East Victoria, Australia. In 2003, he left to focus on wine. With partner and former medical researcher Maree Collis and pharmaceutical executive Adrian Thomas, Nadeson has created a 3,000-case winery in Geelong. Lethbridge produces organic, biodynamic wines with a European accent, from varieties including Pinot Noir, Sangiovese and Chardonnay. Nadeson constructed the naturally ventilated buildings of the winery itself from straw bales. Lethbridge’s 2008 Mietta Pinot Noir, according to wine writer Jamie



Goode, has a “taut, fine, spicy, aromatic nose” and a “rich, concentrated, fine-grained palate ... with good acidity and beautiful focus”.





boon. For instance, English wine, once a curiosity, is now being taken seriously: as of 2011, there were 419 vineyards, 124 wineries and 1,384 hectares of vines planted. Famed German winegrower Helmut Dönnhoff feels that the degree rise in average temperatures since 1988 has also been kind to German vignerons, saying that the “acidity is lower and the grapes are riper, so the quality of wines has improved”. In the past, German Rieslings from cooler vintages often struggled to attain ripeness, and ended up with searing levels of acidity.

Climate change isn't a simple story, however. A study published in 2010 showed that the climate in California's Napa Valley, which is strongly influenced by air currents from the Pacific, may even be getting cooler.

But it is Jones' predictions for the coming decades that have winegrowers worried. These suggest an average increase in growing-season temperature of 2.04 °C between 2000 and 2049, with the largest predicted change in southern Portugal (2.85 °C) and the lowest in South Africa (0.88 °C). For many, this could be disastrous. Water availability for irrigation is also a concern.

Rising temperatures also bring an increased frequency of extreme weather events, and climate unpredictability. A vintage can be ruined by hail, frost or rain at the wrong time. In Victoria, Australia, shortly before harvest in early February 2009, three days of unprecedented high temperatures (43–45 °C) were followed by ‘Black Saturday’

on 7 February, when temperatures of 47 °C led to bush fires that killed 173 people. The damage to the region's wine industry caused by fires, smoke taint of grapes and loss of wine quality was estimated at some AU\$368 million.

In some cases, the careful matching of grape variety to vineyard site may have to be reconsidered. This is not a simple solution: grape vines take at least 3 years to produce a crop, and only begin producing peak-quality grapes after 10 or 15 years. And, in most European regions, growers are allowed to plant only authorized varieties.

Now there is a wind from the east, too. China is becoming a leading consumer of wine — and is already the world's eighth-biggest producer. Will parched Europeans, Australians, Americans and Chileans be drinking Great Wall Wine from Hebei a decade down the pike? The warming wine world is watching with interest. ■

Jamie Goode is the author of *The Science of Wine* and publisher of www.wineanorak.com.
e-mail: jamie@wineanorak.com

Ridge, Santa Cruz Mountains, California

One of California's most famous wineries, Ridge was launched in 1959 when four scientists at the Stanford Research Institute (SRI) reopened the nineteenth-century Monte Bello winery. Those four were engineer Howard Zeidler; David Bennion, who worked with magnetic systems; Charles Rosen, long-time head of SRI's Robotics Division; and pioneering computer engineer Hewitt Crane, who had collaborated with John von Neumann at the Institute for Advanced Study in Princeton, New Jersey. Bennion made the wines until 1969, when philosopher Paul Draper took over. In the 1970s, Ridge showed the world the depth and complexity possible in Californian Cabernet Sauvignon. The 2009 Monte Bello — the winery's flagship wine — is characterized by Draper as having a “blackberry, red currant, wet stone” nose and a richly textured palate with “minerality, elegant chalky structure and firm acidity”.



Domaine de Vens-le-Haut, Savoie, France

Georges Siegenthaler is on a mission to develop micro-cuvées — limited-production, exceptional wines, made as if in a laboratory. After a career as a biochemist at the University of Geneva's faculty of medicine, Siegenthaler established Domaine de Vens-le-Haut (DVH) in France's Rhône-Alpes region of Savoie in 2002. His watchword is that good wine is made in the vineyard, not the cellar. So although he is serious about alternative methods of fermentation and flavours, ‘clean’ viticulture is his passion. In particular, he and his team are working on alternative treatments for common vine mildews; the conventional one, copper sulphate, accumulates in soil in doses that are toxic to wildlife. Siegenthaler notes that the DVH Mondeuse noire, the great red varietal of the region, is dark ruby in colour “with black berry fruits ... and spicy notes”, a dense structure and “gentle tannins in the finish”.

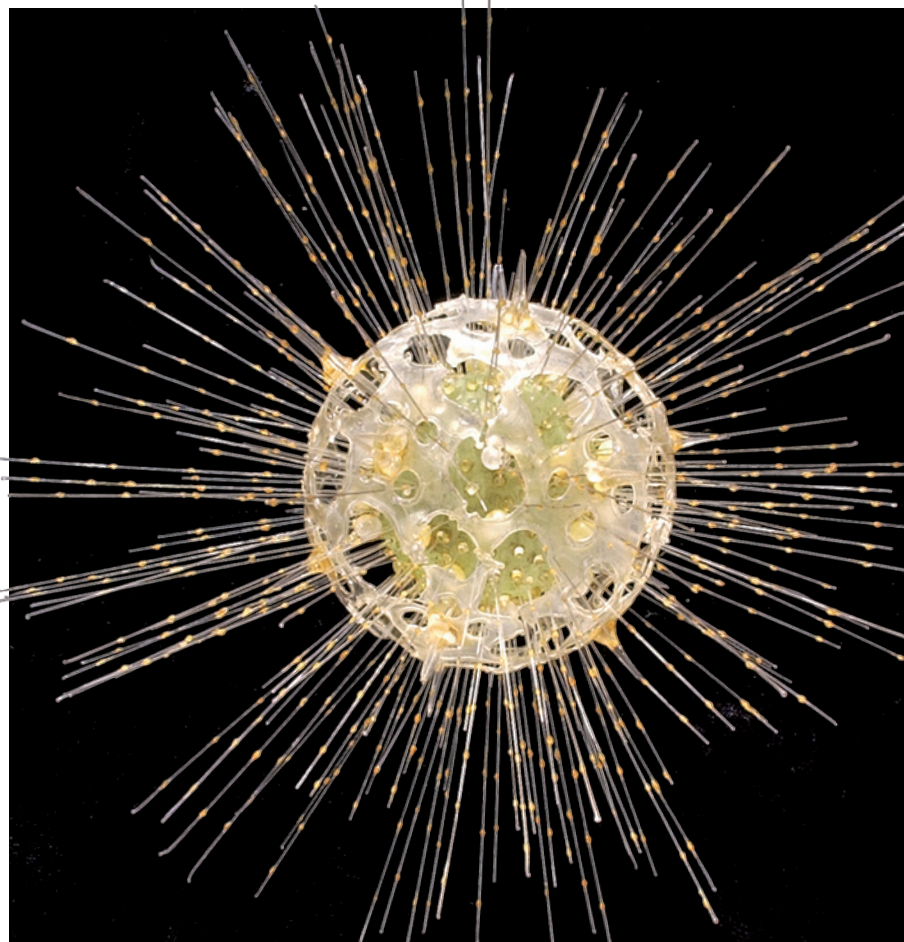


Domodimonti Società Agricola, Le Marche, Italy

Organic chemist Francesco Bellini — who helped to develop the first compound anti-HIV drug to be made commercially, Epiriv (3TC) — is a self-made man. A mid-1960s migrant to Canada from the Marche region of Italy, he has since authored or co-authored more than two dozen patents. After setting up pharmaceuticals company BioChem Pharma, Bellini began to build Domodimonti Società Agricola, a winery in Montefiore dell'Aso, Marche, in 2004. Bellini sees wine-making as the most ancient form of biochemistry. His ‘natural wines’ are based on sustainable practice, such as hand-picking the grapes and banning additives that enhance mouth-feel or colour. Domodimonti's 2008 Il Messia, made from Montepulciano and Merlot, is a deep ruby red with real biochemical verve — fruity, “powerful and fleshy”, as the winery puts it, with “emphyreumatic notes and anise”.



Research by Jamie Goode, wine writer



A glass model of a radiolarian created by Leopold and Rudolf Blaschka in the late nineteenth century.

NATURAL HISTORY

Diamonds in the rough

Ewen Callaway delights in a cherry-picked selection of the London Natural History Museum's gargantuan trove.

Thousands thronged London's Tate Modern gallery this spring to see a diamond-encrusted platinum cast of a human skull worth a cool £50 million: Damien Hirst's *For the Love of God*. Now, across town at the venerable Natural History Museum, gems of a different sort await.

The *Treasures* exhibition features another skeletal presence, an *Archaeopteryx* imprisoned in a slate slab, along with 21 other well-chosen objects. Not convinced that the 150-million-year-old slab competes with Hirstian bling? True, the museum paid German collector Karl Häberlein just £700 (equivalent to about £50,000 today) for the fossil in the 1860s. But *Archaeopteryx* offered the first evidence that birds evolved from dinosaurs, and palaeontologists considered the beast to have been the first bird until only recently, when fossils from China suggested it

Treasures

NATURAL HISTORY MUSEUM, LONDON

might instead have been a feathered dinosaur.

The fossil, which is on display for the first time, was last year declared the 'type specimen' of *Archaeopteryx*, and is the only one with a fully preserved brain imprint on the inside of its skull. Such anecdotes are a highlight of *Treasures*, and are engagingly told using interactive touch displays besides each piece. They provide necessary context for the curious non-scientist, and show how natural history museums can remain relevant in a world of cheap travel and ubiquitous high-definition footage.

Few objects are more fetching than a glass rendering of the single-celled *Aulosphaera elegantissima* — a member of the radiolarian group of protozoa — enlarged tens of times.

It is one of 185 models that the museum purchased in the late nineteenth century from the Blaschkas, a glass-working family based in Germany.

It is a shame that there is room for only one page from the world's most expensive book, John J. Audubon's extraordinary *The Birds of America*. The pages were released serially between 1827 and 1838, and with only 120 complete original editions of the metre-high book known to survive (two owned by the Natural History Museum), copies now sell for millions of dollars. Fortunately, however, there is an interactive version at the free exhibition's entrance, allowing visitors to flick through the natural-history classic, renowned for both action-packed depictions and Romantic styling.

Unsurprisingly, extinction abounds here. The twin bird exhibits — a dodo skeleton and a preserved great auk — represent creatures that humans hunted to death, and speak of how natural history and human history interweave.

Scientists disagree over our role in the extinction of Neanderthals, *Homo sapiens'* big-bodied relative, which died out around 30,000 years ago. The species is named after Germany's Neander Valley, where the first recognized Neanderthal fossilized remains were found in 1856. If the skull on display at the *Treasures* exhibition had been correctly classified when it was discovered eight years earlier, we would be speaking of Gibraltar Man. The exhibit explains that scientists collected DNA from other Neanderthals and discovered that the species interbred with humans, reinforcing the fact that dinosaur-packed institutions such as the Natural History Museum also host cutting-edge science.

Yet another theme of *Treasures* is the history of natural history. An example is provided by three exhibits displayed side by side: a signed first edition of Charles Darwin's *On the Origin of Species*; the remains of pigeons that helped Darwin to make a case for natural selection; and a stunning lepidoptera collection that once belonged to Alfred Russell Wallace, the overlooked co-proponent of evolution by natural selection.

Also on display are some of the 200-million-year-old ammonite fossils that led William Smith, the father of modern geology, to deduce that Earth consists of geological layers that record the deep past. And herbarium pressings, belonging to a director of the Dutch East India Company, George Clifford, that helped a young doctor named Carl Linnaeus to devise a revolutionary naming system for all species.

Rocks, bones, dried plants and insects may never fetch as much on the black market as diamonds and precious metals. But the ideas they inspired are priceless. ■

Ewen Callaway is a reporter for Nature.

NUTRITION

Sugar caned

David Katz finds much to chew on in a polemic on the risk of consuming too much high-fructose corn syrup.

Excess sugar deserves to be in the spotlight that Robert Lustig shines on it. And *Fat Chance*, at its best, is genuinely illuminating.

Lustig, a medical doctor and endocrinologist, has provided obesity treatment to children and adolescents for 16 years — during which time the condition has overtaken nearly 20% of that age group in the United States. In 2009, his YouTube video *Sugar: The Bitter Truth* went viral. Now, in *Fat Chance*, he lays out the case that sugar — and fructose in particular — is the leading cause of obesity and metabolic diseases in both young people and adults. He also touches on the importance of other dietary factors, such as fibre.

The book is, ostensibly, Lustig's opportunity to explain in detail why fructose — present in many foods in syrup derived from heavily subsidized maize (corn) and other sources — is the one thing really wrong with our diets and health. But it is also an opportunity to acknowledge that there is more to the story. He manages the reconciliation of these two conflicting agendas with variable success.

Some solid science underlies *Fat Chance*. Lustig has published dozens of articles laying out his case against sugar in general, and fructose in particular. He participated in the development of an official position on the role of sugar in cardiovascular health for the American Heart Association. His book is at its best when telling us why fructose is bad.

Dietary sugars trigger a release of the hormone insulin, which facilitates both the uptake of glucose by our cells for immediate use as fuel and the conversion of excess calories into fat, for storage. Lustig makes the case that fructose is preferentially stored as fat, and preferentially taken up by the liver, where fat accumulation is most harmful. Fructose, Lustig notes, shares metabolic pathways with ethanol, and its consumption is linked to a range of adverse conditions. One is hepatic insulin resistance, in which insulin receptors in liver cells become desensitized to the actions of insulin, leading to a demand for higher levels of the hormone and the risk of pancreatic stress and diabetes. Another, metabolic syndrome, is a constellation of factors associated with insulin resistance, including abdominal obesity and high blood pressure.



Fat Chance:
Beating the Odds
Against Sugar,
Processed Food,
Obesity, and
Disease

ROBERT H. LUSTIG
Hudson Street: 2012.
336 pp. \$25.95

against 'nanny state' intrusions are entertaining and enlightening. The examples — notably one in which a sweetened yogurt is shown to be a dairy product acting as delivery vehicle for a soft drink — can be quite stunning.

But there are inconsistencies. Lustig says that fructose is a particularly bad actor because it does not trigger an insulin release. That lack of insulin, he says, means that levels of the hormone leptin — which influences appetite and fat accumulation — do not rise, leading to persistent hunger and overeating. But elsewhere in the book, Lustig writes that the insulin resistance caused by fructose consumption leads to elevated levels of insulin in the bloodstream, interfering with leptin signalling, and leading to overeating. Can a lack of insulin release and excessive insulin release be the same problem?

And what about glucose? This is a component of most sugars we consume, from sucrose (table sugar) to high-fructose corn syrup to the lactose in milk (which contains equal amounts of glucose and galactose). Glucose does trigger an insulin release, so if high levels of insulin are among the perils of excess sugar intake, the argument that fructose is uniquely bad is suspect. This is compounded by the fact that fruit is the one common source of pure fructose in our diets

— yet even Lustig identifies fresh fruit as part of the solution. The problem is surely eating too much sugar per se, and the consumption of sources such as syrups and juices that contain little or no fibre, for instance.

There are few overt errors in *Fat Chance* — but the assertion that “burning a pound of fat liberates 2,500 calories” is one. The figure is 4,086 calories (454 grams in a pound, and 9 calories per gram of fat). Lustig then goes on to develop a whole argument against the relevance of calories and the role of exercise in weight control based on the error. However, the best data source we have for sustained weight loss, the US National Weight Control Registry, indicates that routine exercise is an almost universal feature in such success.

Fructose, writes Lustig, is Darth Vader. That presumably makes Lustig Luke Skywalker. There is much more to the universe of our dietary concerns than this one encounter, such as unwanted chemicals in food, refined starch and trans fats. But we would surely benefit from eating less sugar, and Lustig is on a mission to make it happen. May the force be with him. ■

David Katz is the founding director of Yale University Prevention Research Center in Derby, Connecticut, editor-in-chief of the journal *Childhood Obesity and president-elect of the American College of Lifestyle Medicine.*
e-mail: david.katz@yale.edu

Correspondence

English not Latin in botanical reports

The decision at last year's International Botanical Congress to allow the use of either Latin or English for botanical descriptions and diagnoses was hailed as a triumph by the community. The ruling came into effect from 1 January 2012, but many botanists persist in using Latin.

There may be an element of understandable cultural resistance to changing to English. But most scientific journals are English-language publications, and there would seem to be no scientific reason for using Latin — it does not aid clarity or accuracy. It may have been effective for communication centuries ago, but now it has gone the way of the dodo.

The huge amount of literature that contains first-species descriptions in Latin poses a big challenge for translators, but that is no reason to compound the problem by continuing to use this ancient language. Instead, linguaphiles could deploy their skills by editing the computer-generated translations used for online botanical descriptions and books, creating a legacy rather than a liability.

English-language journals could contribute by asking for botanical diagnoses (explanations of how new taxa differ from their relatives) and descriptions to be in English not Latin. Botanists would then be freed up to study the world's flora and help land and natural-resource managers to make informed decisions.

Frank Udovicic *National Herbarium of Victoria, Royal Botanic Gardens Melbourne, South Yarra, Victoria, Australia.*
frank.udovicic@rbg.vic.gov.au

Thousand-citation papers are outliers

Joshua Nicholson and John Ioannidis suggest that the US National Institutes of Health (NIH) is failing to fund the most



impactful researchers in favour of more conformist scientists with interests similar to those of the grant reviewers (*Nature* **492**, 34–36; 2012). For presumably practical reasons, Nicholson and Ioannidis define these high-impact scientists as authors of a paper(s) with at least 1,000 citations. But setting the bar so high could yield unrepresentative outliers in the analysis.

Many of these authors are unlikely to repeat this level of impact consistently, so their papers become anecdotal rarities and atypical of career performance. And as the Comment revealed, several authors of papers with 1,000 or more citations had varied and justifiable explanations for not holding NIH funding. Also, agencies such as the Howard Hughes Medical Institute skim off some of the high-performing researchers, effectively removing them from the NIH system because they are considered to be well funded.

There can be network or familiarity bias among grant reviewers towards certain applicants, often for legitimate reasons — such as when an applicant's proposal might benefit from wider collaboration. Indeed, a junior investigator could find it productive to participate in a grant-panel study section, in which he or she can be exposed

to new ideas and learn from peers.

James Woodgett Samuel
Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada.
woodgett@lunenfeld.ca

Invest proceeds of 4G sale in UK science

Earlier this month the UK government pledged £600 million (US\$968 million) to science projects in recognition of the country's need to develop a more high-tech economy to promote sustainable growth. The Campaign for Science and Engineering (CaSE) and the UK charity Nesta are urging the government to supplement this investment with the £4 billion it is shortly due to receive from the sale of the 4G radio spectrum, which will revolutionize broadband speeds for mobile devices.

Our 4Growth public petition to reinvest this windfall in science and technology (see its4growth.co.uk and go.nature.com/kft87z) has already collected almost 2,000 signatures, and has attracted widespread support from politicians and scientists.

To drive technological progress, 4Growth makes detailed recommendations for targeted investment in

people and skills, research infrastructure and commercial scientific enterprises.

This would be a handsome return on the efforts of innovators such as James Clerk Maxwell, Tim Berners-Lee, Guglielmo Marconi and John Logie Baird, without whom we wouldn't even be using 4G technology — let alone raising money from it.

Imran Khan *Campaign for Science and Engineering, London, UK.*

imran@sciencecampaign.org.uk
Stian Westlake *Nesta, London, UK.*

Cite links to data in reference lists

A huge amount of work goes into creating data sets. It is crucial that these data, big or small, should be more prominently linked to their associated research articles as standard practice.

To achieve this, data can be cited directly in a publication's reference section using a permanent identifier such as a digital object identifier (DOI; see, for example, go.nature.com/vnyidi and go.nature.com/zdfbcl). So far, however, only very few journals do this.

Publishers, funders, researchers and institutions all need to recognize that data sets constitute a valuable scholarly resource. Authors should be credited for these career-making contributions. Enhanced data-set visibility would also benefit referees and readers by raising standards of data analysis, promoting more detailed review, encouraging data curation and boosting reproducibility and data reuse.

Laurie Goodman *GigaScience, BGI, Hong Kong, China.*

Rebecca Lawrence *F1000 Research, Faculty of 1000, London, UK.*

rebecca.lawrence@f1000.com

Kevin Ashley *Digital Curation Centre, University of Edinburgh, UK.*

Thermal Physics

Quantum interference heats up

A thermal effect predicted more than 40 years ago was nearly forgotten, while a related phenomenon stole the limelight. Now experimentally verified, the effect could spur the development of heat-controlling devices. [SEE LETTER P.401](#)

RAYMOND W. SIMMONDS

Wouldn't it be strange to have a material whose thermal conductivity could be changed by a magnetic field? Imagine holding the end of a rod made of this material with the other end placed in a hot fire. As long as a friend keeps a bar magnet away from the rod, you wouldn't burn your hand, but as soon as they apply a magnetic field — ouch! As odd as this seems, the rules of quantum mechanics predict this type of situation for heat transported across a pair of Josephson junctions (devices that consist of two superconductors separated by a thin insulating gap). Writing on page 401, Giazotto and Martínez-Pérez¹ report experiments confirming that this strange phenomenon can actually occur.

In 1962, Brian Josephson made a remarkable discovery² as a graduate student, while investigating what would happen if two superconducting metals were placed very close together without touching. He found that the 'Cooper pairs' of electrons that make up the supercurrent (a current that flows without resistance) in superconductors could miraculously jump, or 'tunnel', across the gap without needing an applied electric voltage.

The size of the supercurrent flowing through this 'tunnel barrier' depends on whether the superconductors at either edge of the gap have the same or a different phase — a property of the quantum-mechanical wavefunction that describes the behaviour of Cooper pairs. In a bulk superconductor, any phase changes in the wavefunction between local regions gives rise to supercurrent flow. Alternatively, forcing a supercurrent to flow produces phase differences, even across a thin non-conducting or insulating barrier.

Consider also what happens when superconductors form closed circuits, such as loops. Now the total phase that accumulates around the loop when supercurrent flows must be an integer multiple of 2π , to maintain the continuity of the wavefunction. This causes magnetic flux in the system to be quantized. The Josephson effect can be combined with this flux quantization to produce a superconducting direct-current quantum interference device³ (d.c.-SQUID). In these devices, a split

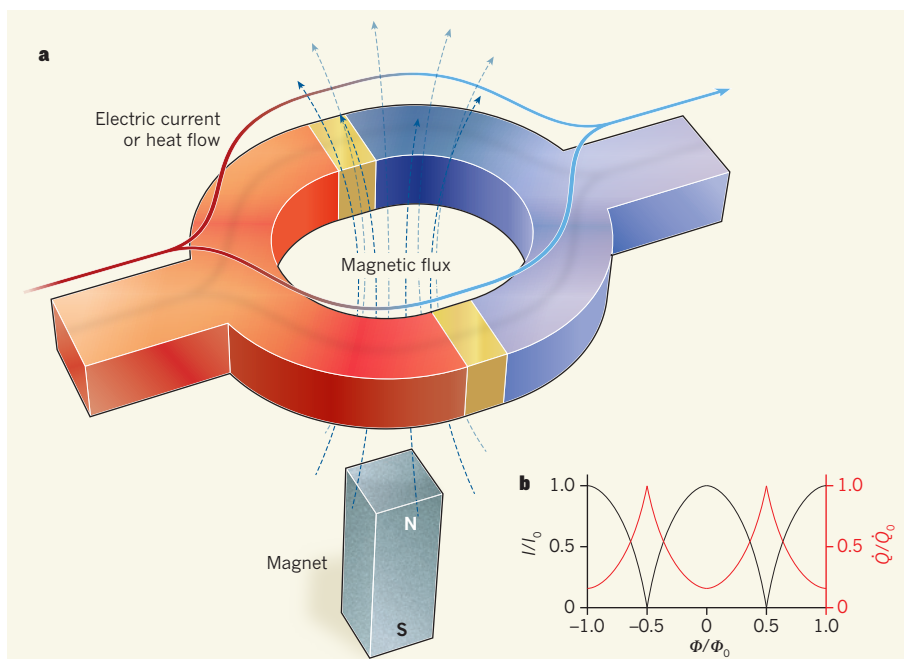


Figure 1 | A direct-current superconducting quantum interference device (d.c.-SQUID). **a**, In d.c.-SQUIDS, a superconducting loop contains two Josephson junctions — thin insulating barriers (yellow) sandwiched between the two superconductors (red and blue). **b**, The maximum electrical current (I , black, left axis) flowing through the device from left to right can be fully modulated by the amount of magnetic flux (Φ) passing through the loop. I_0 is the maximum current that can flow through the d.c.-SQUID; Φ_0 is the magnetic flux quantum, 2.07×10^{-15} webers. Giazotto and Martínez-Pérez¹ have observed an interference effect for heat flow (\dot{Q} , red, right axis; \dot{Q}_0 is the maximum total heat-flow current) through a d.c.-SQUID: the total amount of heat passing through the device can also be modulated by an applied magnetic flux.

superconducting path with two Josephson junctions can sustain a maximum supercurrent, the amplitude of which can be modulated by the amount of magnetic flux piercing the loop (Fig. 1). Such d.c.-SQUIDS are among the most sensitive detectors of magnetic flux ever created and have found many practical applications³.

In addition to the phase-dependent supercurrent, Josephson discovered² two other currents that are present when a finite voltage difference exists across a junction. These currents were caused by the tunnelling of quasiparticles (lone electrons from broken Cooper pairs) or of quasiparticles with Cooper pairs. The first type was similar to the flow of electrons through normal metal-metal junctions, but the second type of current was rather odd: it involved a dynamic

process in which the tunnelling occurred in conjunction with processes for breaking and recombining Cooper pairs. Because Cooper pairs are involved, this current should exhibit interference effects analogous to those seen in d.c.-SQUIDS (in which differences in the wavefunction's accumulated phase along the two paths of a loop create constructive or destructive interference). But electrical experiments that clearly quantify the behaviour of this 'interference current' have remained elusive⁴.

What does all this talk of electrical currents have to do with thermal properties? Well, according to the Wiedemann–Franz law, a metal's thermal conductivity is proportional to its electrical conductivity (and to temperature). This is because electrons can transport some of the heat in a metal. Only three years after

Josephson's seminal work, it was proposed⁵ that thermal conduction through a Josephson junction should involve both quasiparticle flow and the strange interference current that is influenced by the phase across the junction. Most of Josephson's predictions were demonstrated experimentally during the decade following his discovery, but this prediction has lain dormant for many years⁶.

Giazotto and Martínez-Pérez have now observed the interference effect of heat flow in a study that uses essentially the same d.c.-SQUID arrangement as that used for the electrical experiments. By heating one side of the d.c.-SQUID (the red side in Fig. 1) and monitoring the temperature difference across it, the authors verified that the total heat flux through the device can be modulated by an applied magnetic flux, just as predicted⁷. The interference heat current has the interesting property of being able to conduct energy in the opposite direction to the temperature drop across a junction. This allowed the researchers to reduce the quasiparticle heat conductance by tuning the phase difference across each junction using magnetic flux, while maintaining net heat conductance in the direction of the temperature drop, in agreement with the second law of thermodynamics.

The quantum interference of heat currents measured by Giazotto and Martínez-Pérez is unique — it is unlike the standard thermoelectric effects discovered in the early 1800s (in which temperature differences are converted into voltage, and vice versa) and the related, more exotic effects that involve applied magnetic fields. It resembles the interference effects resulting from 'Andreev reflection' within loops formed from junctions between normal metals and superconductors⁸, but those effects are still thermoelectric in nature.

Systems that have interdependent thermal, electric and magnetic properties are immensely important for making practical heat-controlling devices. Exploring superconducting, atomic and molecular junctions, or other nanoscale systems, could lead (and in some cases has already led) to the development of chip-scale heat engines or refrigerators, and energy-harvesting machines⁹. A limitation of superconducting devices is that they need low temperatures to operate, but there are applications for which cryogenic conditions are an advantage. One example is the transition-edge sensor — a device for detecting photons that has made a huge impact on radioastronomy¹⁰. These sensors, and other devices that measure the power of incident electromagnetic radiation, could benefit from Giazotto and Martínez-Pérez's discovery. A device in which thermal conductance can be rapidly tuned using magnetic flux could help to eliminate the heat from a radio telescope's camera pixels, increasing the frame rate for acquiring images and reducing the total time required to map the sky. ■

Raymond W. Simmonds is in the *Quantum Metrology Division, National Institute of Standards and Technology, Boulder, Colorado 80305, USA.*

e-mail: simmonds@boulder.nist.gov

1. Giazotto, F. & Martínez-Pérez, M. J. *Nature* **492**, 401–405 (2012).
2. Josephson, B. D. *Phys. Lett.* **1**, 251–253 (1962).
3. Clarke, J. & Braginski, A. I. (eds) *The SQUID Handbook* (Wiley-VCH, 2004).

4. Langenberg, D. N. *Rev. Phys. Appl.* **9**, 35–40 (1974).
5. Maki, K. & Griffin, A. *Phys. Rev. Lett.* **15**, 921–923 (1965).
6. Guttman, G. D., Nathanson, B., Ben-Jacob, E. & Bergman, D. J. *Phys. Rev. B* **55**, 3849–3855 (1997).
7. Giazotto, F. & Martínez-Pérez, M. J. *Appl. Phys. Lett.* **101**, 102601 (2012).
8. Eom, J., Chien, C.-J. & Chandrasekhar, V. *Phys. Rev. Lett.* **81**, 437–440 (1998).
9. Dubi, Y. & Di Ventra, M. *Rev. Mod. Phys.* **83**, 131–155 (2011).
10. Irwin, K. D. *Sci. Am.* **295**, 86–94 (2006).

NEURODEVELOPMENTAL DISORDERS

Signalling pathways of fragile X syndrome

The RNA-binding protein FMR1 has a key role in the neurodevelopmental disorder fragile X syndrome, but the RNAs targeted by the protein were mostly unknown. An analysis reveals thousands of possible targets. [SEE ARTICLE P.382](#)

SABARINATH JAYASEELAN & SCOTT A. TENENBAUM

Fragile X syndrome was the first genetic disorder found to link RNA regulation to human cognitive function. Of the forms of hereditary autism or autism spectrum disorder associated with a single gene, this is the most common. It results in a range of intellectual disabilities, from mild to severe, and causes sufferers to exhibit certain physical characteristics. The syndrome is caused by an increase in the number of repeats of a short nucleotide sequence located in the X-chromosome gene *Fragile X mental retardation 1* (*FMR1*), which encodes the RNA-binding protein FMRP. On page 382 of this issue, Ascano *et al.*¹ report the discovery of RNA sequences known as RNA-recognition elements that are targeted by the two independent RNA-binding regions of FMRP — a finding that increases our understanding of the molecular basis of fragile X syndrome*. The authors also describe the binding sites in messenger RNAs targeted by the FMRP family of proteins, and their use of a mutant FMRP to further study the protein's role in translation².

Because FMRP is associated with polyribosomes (clusters of ribosomes, the molecular machines that synthesize proteins) and neuron-specific mRNA, it is thought to have an essential role in the post-transcriptional regulation of gene expression in neurons. Accordingly, FMRP and its associated mRNA targets have been studied intensively for more than a decade. Proteins of the FMRP family have been shown to shuttle between the cell's nucleus and the cytoplasm³, to activate

translation⁴ and, in coordination with the Argonaute 2 protein, to associate with certain elements of tumour-necrosis factor α (a protein that regulates immune cells). In addition, FMRP acts as a translational silencer⁵, especially of mRNAs that encode proteins in synapses⁶ (the junctions between neurons). A 'kissing complex' structure in mRNA targets can interact with one of the binding domains in FMRP, and this may account for the protein's role in mRNA translation⁷. FMRP also has a role⁸ in preventing ribosomal translocation on mRNA.

Fragile-X research has the potential to be a model of translational neuroscience, but the syndrome cannot be properly understood until the mRNAs targeted by FMRP have been identified⁹. Until now, no specific RNA-recognition element had been determined for FMRP, and only a handful of the protein's mRNA targets had been confirmed.

By using a combination of PAR-CLIP and RIP-Chip — two genomic methods for capturing mRNA-protein complexes and identifying the RNA in them^{10–13} — Ascano *et al.* have shed light on FMRP's target mRNAs and established a correlation between mRNA targets from different cell types. The authors identified hundreds of probable mRNA targets, many of which have previously been implicated in autism spectrum disorder.

Ascano and colleagues' computational analysis of their results revealed two main

“The main strength of this study lies in its unbiased search of the miRNAs, and thus of the total cellular pool of proteins that they could have affected.”

*This article and the paper under discussion¹ were published online on 12 December 2012.

RNA-recognition elements that are bound by FMRP and found in most of the mRNAs. Further analysis also showed that FMRP regulates the production of two proteins: mTor (which forms part of a signalling pathway related to cell growth and proliferation) and Tsc2, both of which have a role in cancer progression.

Fragile X syndrome is largely considered to be a neurological disorder, but other organs, such as the ovaries, can also be affected; there has been less research into these non-neurological aspects, however. The authors therefore validated their PAR-CLIP and RIP-Chip results by examining the ovaries of mice that had been genetically manipulated to knock out the *Fmr1* gene. They observed that the ovaries of these mice are larger and have higher levels of fragile-X-regulated proteins than the ovaries of normal mice. They also found that concentrations of proteins related to fragile X syndrome were higher in post-mortem brain samples from humans with the disorder than from those without it.

Although Ascano and colleagues' study identifies many of the mRNA targets of FMRP, and points to their probable role in fragile X syndrome and autism spectrum disorder, it is unlikely to be the final word on the matter. Genomic studies of this kind are often criticized, partly because the raw data always suffer from signal-to-noise issues, but also because some researchers question the analytical methods used. Such studies also typically identify hundreds, if not thousands, of mRNAs — it will be a huge task to perform the in-depth analysis required to determine the effects of FMRP binding to each of these targets.

Another potential problem is that Ascano *et al.* performed their RNA-sequencing experiments in human embryonic cell lines, rather than in neuronal cells. Although there is some genetic overlap between these cell types, it will be important to determine the effects of varying FMRP levels in neurons, and this is technically difficult. The role, if any, of microRNAs — short RNAs that modulate a variety of biological functions — in the regulation of fragile-X signalling pathways is also unknown, because PAR-CLIP and RIP-Chip have a limited ability to identify microRNAs as protein-binding targets. Indeed, FMRP is known to associate with the RNA-induced silencing complex^{14–16} (a protein complex that uses a short RNA strand, such as a microRNA, to target complementary mRNAs for degradation), which implies that microRNA regulation and FMRP function are linked.

Ascano and co-workers' findings advance our understanding of the molecular basis of the defects that lead to fragile X syndrome and autism spectrum disorder. In so doing, they might open up avenues of research for drug discovery targeting these neurological disorders, for which there are presently no effective treatments. But it will probably also fan the

flames of controversy surrounding the causes of these conditions. ■

Sabarinath Jayaseelan and Scott A. Tenenbaum are at the College of Nanoscale Science and Engineering, Nanobioscience Constellation, University at Albany, State University of New York, Albany, New York 12203, USA.

e-mails: stenenbaum@albany.edu; sjayaseelan@albany.edu

1. Ascano, M. Jr *et al.* *Nature* **492**, 382–386 (2012).
2. Feng, Y. *et al.* *Mol. Cell* **1**, 109–118 (1997).
3. Eberhart, D. E. & Warren, S. T. *Somat. Cell Mol. Genet.* **22**, 435–441 (1996).
4. Vasudevan, S. & Steitz, J. A. *Cell* **128**, 1105–1118 (2007).

5. Lagerbauer, B., Ostareck, D., Keidel, E. M., Ostareck-Lederer, A. & Fischer, U. *Hum. Mol. Genet.* **10**, 329–338 (2001).
6. Bear, M. F., Huber, K. M. & Warren, S. T. *Trends Neurosci.* **27**, 370–377 (2004).
7. Darnell, J. C. *et al.* *Genes Dev.* **19**, 903–918 (2005).
8. Darnell, J. C. *et al.* *Cell* **146**, 247–261 (2011).
9. Bassell, G. J. & Warren, S. T. *Neuron* **60**, 201–214 (2008).
10. Hafner, M. *et al.* *J. Vis. Exp.* **41**, e2034 (2010).
11. Jayaseelan, S., Doyle, F., Currenti, S. & Tenenbaum, S. A. *Meth. Mol. Biol.* **714**, 407–422 (2011).
12. Keene, J. D., Komisarow, J. M. & Friedersdorf, M. B. *Nature Protoc.* **1**, 302–307 (2006).
13. Tenenbaum, S. A., Lager, P. J., Carson, C. C. & Keene, J. D. *Methods* **26**, 191–198 (2002).
14. Caudy, A. A., Myers, M., Hannon, G. J. & Hammond, S. M. *Genes Dev.* **16**, 2491–2496 (2002).
15. Ishizuka, A., Siomi, M. C. & Siomi, H. *Genes Dev.* **16**, 2497–2508 (2002).
16. Jin, P. *et al.* *Nature Neurosci.* **7**, 113–117 (2004).

CARDIOVASCULAR BIOLOGY

A boost for heart regeneration

Heart muscle cells die en masse after injury, yet the adult mammalian heart retains little capacity to regenerate them. Regulatory microRNA sequences may stimulate self-renewal of these muscle cells. [SEE ARTICLE P.376](#)

MARK MERCOLA

Muscle cells of the mammalian heart exit the cell-division cycle shortly after birth and rarely divide again. Not surprisingly, therefore, after a heart attack — which can kill upwards of a billion of these cardiomyocytes — heart function is compromised, and this can lead to chronic heart failure or even sudden death. This dismal picture has prompted a long-standing search for ways to mend the damaged heart, with particular emphasis on understanding the reasons why cardiomyocytes do not re-enter the cell cycle. Eulalio *et al.*¹ report on page 376 of this issue that certain microRNA sequences can stimulate division of adult cardiomyocytes. Their observations hint at the possibility of identifying the molecular regulators that maintain cardiomyocytes in the quiescent state*.

Over the past 30 years, the prevailing view that heart muscle cells are not replaced has gradually given way to one that there is a modest capacity for cardiomyocyte renewal, albeit outmatched by the loss of these cells after cardiac ischaemia (limited blood supply, and thus oxygen, to the heart) or other types of injury. Perhaps the most compelling evidence for cardiomyocyte regeneration comes from analysis of the incorporation of radioactive carbon-14 into human DNA that occurred worldwide

during the era of above-ground testing of atomic weapons. These data² showed that some 1% of cardiomyocytes in young adults are replaced annually.

Although the incidence of cardiomyocyte turnover declines with age, it might nonetheless contribute to repair after injury. What's more, in mice^{3,4} and humans⁵, a heart attack (myocardial infarction) seems to increase cardiomyocyte turnover within the preserved area of the heart bordering the region of damage, supporting the view that regeneration is an adaptive response to injury.

New muscle cells can come from two sources — from *de novo* differentiation of stem cells and from replication of pre-existing cardiomyocytes. Studies that tracked the fate of existing cardiomyocytes after injury in genetically engineered mice did not detect their replication, attributing renewal to *de novo* differentiation instead³. However, a study⁶ based on a sensitive, mass-spectroscopy technique has confirmed that a low level of replication occurs. Moreover, a number of soluble factors (including neuregulin⁷, periostin⁸ and fibroblast growth factor plus p38 inhibitors⁹) stimulate replication, thereby preserving

cardiac function after ischaemic injury. This leads to the question of what normally prevents cardiomyocyte division.

► NATURE.COM

For more on heart-cell self-renewal, see: go.nature.com/z5fl3r

*This article and the paper under discussion¹ were published online on 5 December 2012.

To address this, Eulalio *et al.* conducted an unbiased search for microRNAs (miRNAs) that can induce re-entry of postnatal cardiomyocytes into the cell cycle. MicroRNAs are natural, single-stranded RNAs that are around 22 nucleotides long and bind directly to and suppress messenger RNA targets through imprecise base-pairing. Active miRNAs thus suppress expression of cellular proteins from mRNAs, and so their identification can provide insight into the protein regulators that keep cells in a quiescent state.

The authors' screen yielded 204 miRNAs that more than doubled the rate of division of postnatal rat cardiomyocytes (and more than 300 miRNAs that suppressed division). Of the 204, a minority (40) also functioned in mice, and Eulalio *et al.* selected two of these (dubbed hsa-miR-590-3p and hsa-miR-199a-3p) for further analysis. In the follow-up studies, they gathered impressive microscopy images to document mitotic cell division and even its final step, cytokinesis.

The researchers also tested the effect of the two miRNAs *in vivo*. When delivered directly to the rat heart — by means of a viral vector that is used in gene therapy — these miRNAs increased cardiomyocyte proliferation. Moreover, the two miRNAs boosted the normally ineffective process of heart repair after infarction in mice. Specifically, several parameters of cardiac function and structure improved over time, despite the initial loss of cardiac muscle due to the infarction.

These findings are notable because they not only add to evidence for the replicative capacity of cardiomyocytes, but also show that miRNAs can modulate cell regeneration in the heart. Furthermore, by overriding the cellular processes that maintain the quiescent state, they open the door to a broader, systems-level understanding of what prevents cardiomyocytes from re-entering the cell cycle and dividing. This information might enable the design of drugs to target some of the proteins whose expression is affected by these miRNAs, enhancing regeneration.

The miRNAs that Eulalio *et al.* tested did not affect the division of another class of cell in the heart, cardiac fibroblasts, suggesting that these regulatory sequences tap into fundamental mechanism(s) of cardiomyocyte differentiation and quiescence. It remains to be discovered whether the same miRNAs also mediate the regrowth of blood vessels in ischaemic tissue and/or blunt adverse changes that occur upon prolonged stress, such as a detrimental shift in energy utilization or a decline in contractility of the heart muscle. Moreover, it should be investigated whether boosting cardiomyocyte proliferation using miRNAs has undesirable consequences that might compromise heart function in the long term.

In my opinion, the main strength of this study lies in its unbiased search of the miRNAs, and thus of the total cellular pool of proteins

(the proteome) that they could have affected, in cardiomyocytes. Functional screening using synthetic miRNAs has also been performed previously to study heart development¹⁰ and disease-related increase in cardiomyocyte size (hypertrophy)¹¹. Given that the 1,500 or so known human miRNAs control an estimated 60% of the proteome¹², this approach is emerging as an effective means of gaining comprehensive insight into complex biology. Eulalio *et al.* used a library of about 900 synthetic miRNAs, thus probing a large swathe of the proteome. The miRNA hits that they got might not be the natural regulators of cardiomyocyte quiescence (using miRNAs with an antisense sequence in a similar screening is more likely to lead to miRNAs that sustain quiescence), but are nevertheless valuable because of the proteins they downregulate.

Unfortunately, identifying the target proteins of miRNAs is not easy. The binding of an miRNA to its cognate mRNA is not just a matter of base-pair recognition: sequences

“The main strength of this study lies in its unbiased search of the miRNAs, and thus of the total cellular pool of proteins that they could have affected.”

surrounding the recognition site and RNA-binding cofactors also influence site accessibility¹². This makes computational approaches too error-prone to use as the sole means of target prediction. Eulalio *et al.* therefore profiled changes in the levels of mRNA gene transcripts in cardiomyocytes treated with hsa-miR-590-3p and hsa-miR-199a-3p. They found about 700 gene transcripts that showed decreased expression and 1,000 transcripts with enhanced expression; there was little overlap between the mRNAs affected by the two miRNAs.

The authors attempted to determine the importance of particular transcripts by reducing their expression using specific short interfering RNAs (siRNAs). Individual siRNAs only partially reproduced the miRNA effects, in agreement with the expectation that the miRNAs act on multiple cellular mRNA targets and, thus, that multiple proteins cooperate to maintain quiescence. Clearly, systems-level analysis of changes to the proteome is needed to further the understanding of cardiomyocyte quiescence.

Knowledge of the mechanisms that keep cardiomyocytes in a differentiated, non-dividing state will greatly aid the development of regenerative drugs. Ideally, such a drug would target an intracellular signalling pathway coordinating several processes that enable a cardiomyocyte to divide successfully, including cell-cycle entry, faithful DNA synthesis to avoid triggering apoptotic cell death, and disassembly and rebuilding of the contractile apparatus. This is a tall order. The present study is important because it takes a step



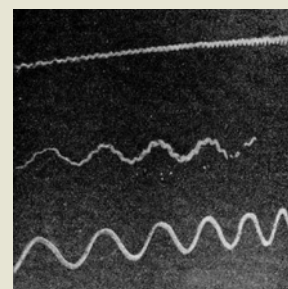
50 Years Ago

I wish to appeal to my fellow geologists and workers in the rapidly growing subject of geochronology to discontinue a habit or fashion which is both unnecessary and misleading. I refer to the habit of calling radiometric ages ‘absolute’ ages. An age does not become ‘absolute’ by virtue of being expressed in units of time such as a year ... The term is not only redundant and both philosophically and scientifically without meaning; it is also misleading in its psychological suggestion of a higher degree of accuracy than can be justified ... If any of the culprits claim that they are in good company with Newton and Kelvin, who both wrote of ‘absolute time’, they have only to remember Einstein and the coming of relativity.

From *Nature* 22 December 1962

100 Years Ago

‘Smoke trace of compound vibrations of tuning-fork’ — The accompanying print is from one corner of a smoke trace used by me at a popular lecture in 1901. One curve shows the fundamental (128 per second), another the first upper



partial, while the centre curve of the three shows the form of vibration executed when the first upper partial is sounding, together with the prime. The three sounds may be heard by the audience, and the smoke traces of each obtained in their presence.

From *Nature* 19 December 1912

towards achieving a systems-level understanding of what cardiomyocyte quiescence means, and addressing whether it can be overcome therapeutically. ■

Mark Mercola is at the Sanford-Burnham Medical Research Institute, La Jolla, and in the Department of Bioengineering, University of California, San Diego, California 92037, USA.

e-mail: mmercola@sanfordburnham.org

1. Eulalio, A. *et al.* *Nature* **492**, 376–381 (2012).
2. Bergmann, O. *et al.* *Science* **324**, 98–102 (2009).
3. Hsieh, P. C. H. *et al.* *Nature Med.* **13**, 970–974 (2007).
4. Olivetti, G. *et al.* *Am. J. Pathol.* **141**, 227–239 (1992).
5. Quaini, F. *et al.* *Circ. Res.* **75**, 1050–1063 (1994).
6. Senyo, S. E. *et al.* *Nature* <http://dx.doi.org/nature11682> (2012).
7. Bersell, K., Arab, S., Haring, B. & Kühn, B. *Cell* **138**, 257–270 (2009).
8. Kühn, B. *et al.* *Nature Med.* **13**, 962–969 (2007).
9. Engel, F. B., Hsieh, P. C. H., Lee, R. T. & Keating, M. T. *Proc. Natl Acad. Sci. USA* **103**, 15546–15551 (2006).
10. Colas, A. *et al.* *Genes Dev.* **26**, <http://dx.doi.org/10.1101/gad.200758.112> (2012).
11. Jentzsch, C. *et al.* *J. Mol. Cell. Cardiol.* **52**, 13–20 (2012).
12. Bartel, D. P. *Cell* **136**, 215–233 (2009).

CANCER

Metabolism in the driver's seat

It is increasingly accepted that metabolic changes in cancer cells can drive tumour formation. The finding that the SIRT6 protein suppresses tumour formation by regulating metabolism adds weight to this view.

LUISA TASSELLI & KATRIN F. CHUA

Cancer development is accompanied by extensive reprogramming of cellular energy metabolism¹. A cancer-associated metabolic switch, called the Warburg effect, activates glycolysis — a biochemical pathway that fuels tumour growth by supporting the high energetic and biosynthetic demands of cancer cells. Growing evidence suggests that flipping the switch back, away from glycolysis, could be an effective strategy for tumour suppression. Writing in *Cell*, Sebastián *et al.*² provide further support for this idea. They show that the protein SIRT6

prevents tumorigenesis by opposing cancer-promoting metabolic programs. Intriguingly, SIRT6 is not a classic metabolic enzyme, but rather a protein that regulates the integrity and proper expression of the genome³. Sebastián and colleagues propose that, through gene regulation, SIRT6 controls a key metabolic node that could be targeted to combat cancer.

Tumour initiation and progression involves multiple steps during which cells acquire several hallmarks of cancer, including resistance to cell death and evasion of growth-suppressive cues, that allow them to bypass normal checks on cell proliferation⁴. Much work on cultured cells has shown that several cellular pathways

must be deregulated⁵ for cells to become tumorigenic. For example, in mouse embryonic fibroblasts (MEFs), inactivation of a single tumour suppressor can cause cellular immortalization — the characteristic ability of cancer cells to divide without limit. But these immortalized cells are not cancerous and do not form tumours when transplanted into mice, unless additional tumour-suppressive or oncogenic pathways are deregulated.

Sebastián *et al.* find that SIRT6 deficiency provides the additional 'hit' required to transform immortalized MEFs into proper cancer cells. They also report that in cancer cells, and in a mouse model of colorectal cancer, SIRT6 deficiency increases tumour formation. Thus, SIRT6 loss may promote both cancer initiation and cancer progression (Fig. 1). These findings fit well with previous discoveries that SIRT6 opposes liver-cancer initiation⁶ and controls multiple potentially tumour-suppressive pathways³.

Analysis of human-tumour databases helped Sebastián and co-authors to validate the idea that SIRT6 is an authentic tumour suppressor. They found frequent deletions of the SIRT6 gene and reductions in its messenger RNA levels in pancreatic and colorectal tumour samples, and showed that

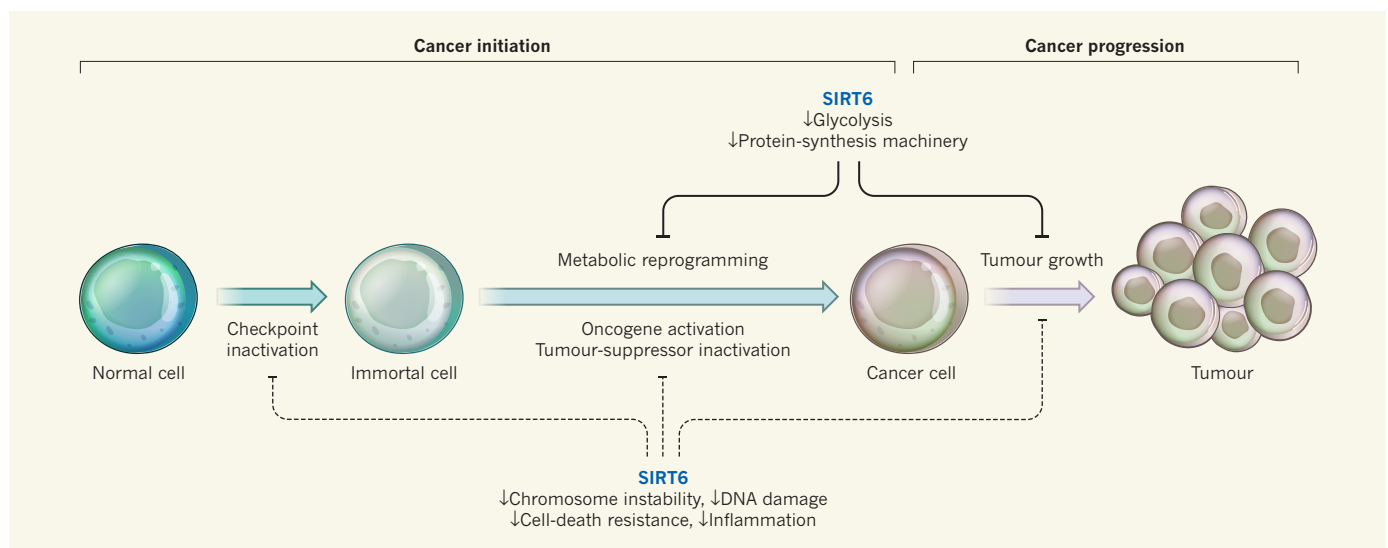


Figure 1 | On the road to cancer. For a normal cell to become cancerous, it must not only become immortal, acquiring the ability to divide indefinitely by escaping cellular checkpoints, but also undergo further oncogenic modifications. The cancer cells can then give rise to tumours. Sebastián *et al.*² report that the protein SIRT6 can prevent

both cancer initiation and cancer progression by suppressing the cancer-associated metabolic processes of glycolysis and protein synthesis. Other previously known roles of SIRT6 suggest that the protein could also contribute to tumour suppression at other stages of cancer development (dashed lines).

low levels of SIRT6 are associated with poor patient outcomes. These findings have clinical implications: SIRT6 expression could serve as a biomarker for cancer prognosis, and SIRT6-activating compounds might be powerful drugs. Indeed, several features of SIRT6 make it an especially promising pharmacological target. In particular, it is a highly selective enzyme that removes acetyl groups from only a few proteins³, and its known cellular functions are all consistent with beneficial effects in the context of cancer. Therapeutic activation of SIRT6 might therefore have fewer unwanted effects than modulating the activity of more promiscuous enzymes.

There is growing evidence that tumour suppressors often control several aspects of cancer biology, and SIRT6 is no exception³ (Fig. 1). It prevents chromosome breakage and abnormal rejoining, and promotes DNA repair — activities that protect against mutations that can fuel tumour progression. It also regulates gene-expression programs that have roles in cancer. For instance, it prevents excessive gene activation by NF- κ B — a transcription factor that promotes cell proliferation, resistance to cell-death signals, and metabolic and inflammatory processes that favour tumour initiation and growth. Sebastián *et al.* report yet another tumour-suppressing effect of SIRT6: it limits the expression of genes that are involved in protein synthesis and which are activated by the MYC oncogene to support tumour growth.

Despite these and other possible tumour-suppressive mechanisms of SIRT6 action, the researchers propose that it is the metabolic rewiring triggered by inactivation of this protein that is essential to making cells cancerous. Indeed, SIRT6 is known to oppose gene activation by HIF1, a transcription factor that activates key enzymes of glycolysis and increases glucose uptake into cells. The authors also show that reversing the increased glycolysis in cancerous SIRT6-deficient cells — by inactivating a central glycolytic enzyme or by reintroducing SIRT6 — reduces the cells' tumorigenic potential.

Sebastián and colleagues further propose that even without any other oncogenic changes, increased glycolytic metabolism due to SIRT6 loss is sufficient for tumour initiation. That a single genetic change is enough to transform normal cells into tumour cells would be a remarkable exception to the rule that combined deregulation of several pathways is required for oncogenic transformation. If proven, this could indicate that metabolic change associated with SIRT6 loss supplies a more powerful oncogenic drive than most other known cancer-promoting mutations.

Alternatively, other effects of SIRT6 loss on tumour-suppressive pathways could favour early steps in cancer development. Future work should ask how various functions of SIRT6 contribute to cancer prevention, and how these intersect with cellular metabolic control.

It should also be determined whether SIRT6 loss causes spontaneous tumour formation in animals. A study⁷ this year evaluated SIRT6-deficient mice up to 17 months of age and, surprisingly, found no evidence of tumour formation. Although the number of mice investigated was small, this observation seems to be at odds with the strong tumorigenic potential of SIRT6-deficient cells that Sebastián *et al.* describe. Studying cancer development in larger cohorts of SIRT6-deficient mice is necessary to resolve this issue.

What molecular mechanisms underlie the observed metabolic changes in SIRT6-deficient cells? Both the HIF1 and MYC pathways have central roles in the metabolic reprogramming of cancer cells, and NF- κ B signalling and genomic instability can also promote glycolytic metabolism^{1,6}. Perhaps SIRT6 prevents metabolic rewiring in cancer through several pathways, which might be redundant or operate in specific cell types or physiological contexts.

Sebastián and collaborators propose that inhibition of glycolytic metabolism could be used to treat tumours that have low SIRT6 levels. Perhaps equally exciting is the prospect that activation of SIRT6 might be beneficial even in tumours that arise independently of changes in

SIRT6. If increasing the amount or activity of SIRT6 can indeed reduce the tumorigenicity of cells transformed by other oncogenic manipulations, strategies involving SIRT6 activation could eventually be widely used for tumour suppression. It is noteworthy that SIRT6 attenuates ageing-related cellular processes, and increasing its activity in mice can extend lifespan^{3,8}. So SIRT6 activation might not only provide an anticancer strategy, but also serve as a preventive medical intervention throughout life. ■

Luisa Tasselli and Katrin F. Chua are in the Division of Endocrinology, Gerontology and Metabolism, Stanford University School of Medicine; and at the VA Palo Alto Health Care System, Geriatric Research, Education and Clinical Center, Palo Alto, California 94304, USA. e-mail: kfcchua@stanford.edu

1. Levine, A. J. & Puzio-Kuter, A. M. *Science* **330**, 1340–1344 (2010).
2. Sebastián, C. *et al.* *Cell* **151**, 1185–1199 (2012).
3. Tennen, R. I. & Chua, K. F. *Trends Biochem. Sci.* **36**, 39–46 (2011).
4. Hanahan, D. & Weinberg, R. A. *Cell* **144**, 646–674 (2011).
5. Hahn, W. C. *et al.* *Nature* **400**, 464–468 (1999).
6. Min, L. *et al.* *Nature Cell Biol.* **14**, 1203–1211 (2012).
7. Xiao, C. *et al.* *J. Biol. Chem.* **287**, 41903–41913 (2012).
8. Kanfi, Y. *et al.* *Nature* **483**, 218–221 (2012).

EPIGENETICS

Erase for a new start

Tet proteins regulate gene expression by removing methyl groups from DNA bases. This activity may be a facilitating step in turning on the cell-division pathway that produces sperm and egg cells. SEE LETTER P.443

SYLVAIN GUIBERT & MICHAEL WEBER

Meiosis is a type of cell division that is a key feature of sexual reproduction. Meiotic division of germ cells produces sperm and egg cells, which have only one copy of each chromosome and which fuse during fertilization to create a new organism with two copies of the genetic material. This highly orchestrated process requires the activation of a specific set of meiotic genes, but how these genes are activated at the right time and in the right place is poorly understood. On page 443 of this issue, Yamaguchi *et al.*¹ report that Tet1, a member of the recently discovered Tet protein family, is required for the activation of meiotic genes in mouse egg cells. This is exciting because Tet proteins are involved in erasing DNA epigenetic marks, suggesting that this is a crucial mechanism in meiosis.

Epigenetic modifications are chemical or structural changes to DNA or DNA-bound proteins that modulate gene expression without changing the DNA sequence. One common epigenetic mark is the addition of a

methyl group to cytosine, one of the four main DNA bases, thereby creating 5-methylcytosine. In mammals, this methylation modifies the accessibility of the DNA molecule and mediates long-lasting silencing of gene expression and of parasitic mobile elements (regions of DNA that can move within the genome)². DNA methylation is essential for the survival of the embryo, and its occurrence is dynamically regulated during development³.

Members of the Tet (ten–eleven translocation) protein family have been implicated in removing this methylation mark from DNA — the proteins can oxidize 5-methylcytosine to form 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine, which are potential intermediates of demethylation^{4–6}. This discovery was groundbreaking, because the mechanisms of DNA demethylation in mammals were a mystery for decades, and it prompted intensive research on the three mammalian Tet proteins (Tet1–3). We now know that Tet3 is expressed in oocytes (egg cells) and is required for the reprogramming of epigenetic marks that occurs during the first

cell divisions of the embryo⁷, and that Tet2 has a key role in blood-cell maintenance and is frequently mutated in human leukaemia⁸. But the biological function of Tet1 has been less clear, because no major developmental defects have been found in Tet1-deficient mice⁹.

Yamaguchi *et al.* studied the physiological function of Tet1 by generating genetically modified mice that do not produce a functional full-length Tet1 protein. They show that the absence of Tet1 leads to decreased fertility and fewer pups in the litters than in normal mice. This finding is consistent with the high expression of Tet1 seen in the precursor cells of eggs and sperm, and was also observed recently in another model of Tet1-deficient mice⁹.

Focusing on females, the authors report that Tet1 deficiency leads to a decrease in the size of the ovaries, an increased incidence of cell death in the ovaries and a reduced number of fully developed oocytes. They also show that the oocytes display meiotic defects, such as a failure to align and segregate chromosome pairs efficiently, together with defective genetic recombination (the 'shuffling' of DNA sequences that occurs during meiotic cell division). This failure to complete meiosis probably explains why a high proportion of developing oocytes in the Tet1-deficient mice are eliminated by apoptotic cell death.

To further study how Tet1 promotes meiosis, the authors examined gene-expression profiles of oocyte-precursor cells from wild-type and Tet1-deficient female mice. Strikingly, they observed that the mutant cells express reduced levels of several meiotic genes, such as those encoding components of the synaptonemal complex, which promotes chromosome alignment and recombination.

The obvious next question was, what is the role of Tet1 in activating these meiotic genes? Germ-cell precursors are known to undergo a genome-wide erasure of DNA methylation during their specification for becoming germ cells³. This affects all types of sequences, including the promoter regions of many meiotic genes that are repressed by DNA methylation in somatic (non-germ) cells and that need to be demethylated in germ cells^{10,11}. (Promoter regions are the DNA sequences at which transcription of a gene is initiated.) Given the potential role of Tet proteins in DNA demethylation, it was tempting to speculate that Tet1 is involved in the demethylation of meiotic genes. To address this, the authors measured DNA methylation in the promoters of three meiotic genes and observed various degrees of residual DNA methylation in Tet1-deficient female germ cells, which could explain the genes' reduced expression.

The authors then investigated whether Tet1 is required for the broader erasure of DNA methylation in germ-cell precursors. They generated genome-scale maps of DNA methylation and, surprisingly, observed that

the absence of Tet1 only marginally impairs genome-wide demethylation. This suggests that Tet1 is required only for demethylation of specific sequences, such as meiotic genes. Unfortunately, the authors' methylation data were at low coverage (meaning that each genomic position was measured only a small number of times), and this prevented detailed analysis of which sequences require Tet1 for demethylation in germ cells. Other aspects that need to be investigated include whether the role of Tet1 in meiotic-gene activation depends solely on its effect on 5-methylcytosine, and whether Tet1 has similar functions in males.

Since the discovery of Tet proteins, their role in epigenetic reprogramming in germ cells has been a matter of speculation. Yamaguchi and colleagues' study provides the first genetic clues about the specifics of this activity, by showing that Tet1 is not essential for general demethylation in germ cells but is required only at certain sequences. The finding raises many questions. Do other Tet proteins compensate for the absence of Tet1? What sequences are demethylated by Tet1,

and how is the protein recruited to these DNA sites? What other mechanisms promote DNA demethylation in germ cells, and how do these processes interact? Could Tet1 be involved in human infertility? We are only beginning to understand the physiological and molecular roles of Tet proteins, and this work adds a new chapter to an exciting story. ■

Sylvain Guibert and Michael Weber are at the UMR 7242 Biotechnology and Cell Signalling, CNRS/University of Strasbourg, 67412 Illkirch, France.
e-mail: michael.weber@unistra.fr

1. Yamaguchi, S. *et al.* *Nature* **492**, 443–447 (2012).
2. Feng, S., Jacobsen, S. E. & Reik, W. *Science* **330**, 622–627 (2010).
3. Saitou, M., Kagiwada, S. & Kurimoto, K. *Development* **139**, 15–31 (2012).
4. Tahiliani, M. *et al.* *Science* **324**, 930–935 (2009).
5. He, Y.-F. *et al.* *Science* **333**, 1303–1307 (2011).
6. Ito, S. *et al.* *Science* **333**, 1300–1303 (2011).
7. Gu, T.-P. *et al.* *Nature* **477**, 606–610 (2011).
8. Quivoron, C. *et al.* *Cancer Cell* **20**, 25–38 (2011).
9. Dawlaty, M. M. *et al.* *Cell Stem Cell* **9**, 166–175 (2011).
10. Borgel, J. *et al.* *Nature Genet.* **42**, 1093–1100 (2010).
11. Hackett, J. A. *et al.* *Development* **139**, 3623–3632 (2012).

LOW-TEMPERATURE PHYSICS

Cool molecules

A sample of the hydroxyl radical has been cooled to a temperature of a few millikelvin. The result opens the door to observing phenomena such as Bose–Einstein condensation of molecules in their ground state. [SEE LETTER P.396](#)

PAUL S. JULIENNE

Since the advent of laser cooling and evaporative cooling more than two decades ago, ultracold trapped atoms at temperatures close to a millionth of a kelvin have been used in a wide range of research involving precision measurement and exotic quantum phenomena. Molecules, on the other hand, have resisted cooling down to even 1 millikelvin. On page 396 of this issue, Stuhl *et al.*¹ describe a breakthrough in molecular cooling. They demonstrate strong evaporative cooling that lowers the temperature from 50 mK to just a few millikelvin and increases the density of a decelerated and trapped sample of the chemically interesting hydroxyl radical (OH).

Why are cold atoms or molecules so interesting? First, cold means slow — the velocity of an ultracold atom or molecule is so low that its location in space becomes spread out and wave-like, in accordance with the Heisenberg uncertainty principle. The de Broglie wavelength, which is a measure of this quantum delocalization, can be many hundreds of nanometres or larger. Furthermore, an atom or molecule can be either a boson or a

fermion, depending on its spin quantum number. The OH molecule is a composite boson, for example, having an integer spin quantum number. When the phase-space density (the number of atoms or molecules per cubic de Broglie wavelength) becomes large enough, approaching unity, the 'quantum degenerate regime' is reached. In this regime, bosons can exhibit exotic quantum phenomena, such as Bose–Einstein condensation, in which all the particles occupy a single wave-like quantum state. Other quantum phenomena are possible with cold fermions, including particle pairing similar to that seen in superconductivity.

Although cold molecules are in many ways similar to cold atoms, their more complex structure offers a richer range of phenomena. For example, for OH, which has an electric dipole moment, there are strong long-range forces between two molecules. These forces make it easy to control or engineer a wider range of condensed-matter quantum effects with molecules than with atoms^{2,3}. Another difference is that molecules can take part in chemical reactions. Fast chemical reactions of molecular fermions can be turned off if they are cold enough, because the Pauli exclusion principle

ensures that identical fermions cannot get close enough together to react. Flipping one nuclear spin to stop the molecules being identical can turn the reaction back on again⁴. The exquisite control of chemical reactants and molecular collision complexes should open up new vistas for chemistry involving ultracold molecules⁵.

Why are molecules much harder to cool than atoms? Laser cooling is generally ineffective because molecules have a more complex internal structure than atoms. Two atoms that are already cold can sometimes be coaxed together using magnetic fields and light to make a molecule in its quantum ground (lowest-energy) state^{6,7}, but only a few species can be created this way. Stark deceleration, which uses a pulsed sequence of electric fields to slow a beam of polar molecules, provides a general technique for a variety of species⁶. However, the temperature of such molecules is still relatively high, about 50 mK, and the density in a molecular trap is very low. What has been missing is a method to cool the molecules further, and to increase their density — especially their phase-space density.

Stuhl *et al.* take advantage of the OH molecule's unusual properties. It has both an electric dipole moment, which means that it can be decelerated in a standard Stark decelerator, and an unpaired electron, which gives it a magnetic moment, so that it can be confined in a magnetic trap. The OH molecule's ground state has the useful property of being split into two quantum levels of nearly the same energy but with different parity quantum numbers. Molecules in the higher energy state are trapped, but those in the lower energy state are expelled from the trap. Evaporative cooling involves a

slow reduction of the trap depth, allowing the warmer trapped molecules to escape while the cooler molecules achieve thermal equilibrium through collisions with one another that do not change their state. For evaporative cooling to work, the rate of such thermalizing collisions must remain high compared with the rate of loss collisions that convert the molecules to the untrapped lower energy state.

Evaporative cooling is not possible for most molecules because the ratio of thermalizing to loss collision rates is not favourable. But for OH and similar molecules, the upper state can survive for many collisions, allowing evaporative cooling to proceed. The theoretical model described by Stuhl *et al.* explains how two molecules in the upper state experience repulsive van der Waals forces when they are far apart, as a consequence of the opposite parity of the two ground-state levels of OH that have nearly the same energy. These repulsive forces ensure that fast thermalizing collisions occur while keeping the molecules far enough apart to prevent loss collisions.

Stuhl *et al.* lowered their effective trap depth by introducing a microwave-frequency 'knife' that converts trapped molecules to untrapped ones on the outer edges of the trap. By changing the microwave frequency, they could cut closer and closer to the centre of the trap, where the coldest molecules accumulate. Their tenfold decrease in temperature, with small molecular loss, implies a dramatic thousand-fold increase in phase-space density. The temperature may be even lower than reported, because in this low-temperature regime the molecules are at the limit at which temperature can be measured with the current apparatus.

This is an extremely promising advance in molecular cooling. The theoretical model suggests that evaporation should be even more effective as the molecules get colder. There is every reason to expect that an improved experimental apparatus could lower the temperature and increase the phase-space density enough to make a Bose–Einstein condensate from OH molecules. If the same degree of cooling can be achieved in OD, using deuterium (D, or ²H), a quantum degenerate gas of fermions could also be observed.

It seems that evaporative cooling will soon join the method of associating two ultracold atoms as a way of making an ultracold molecule. Laser cooling of at least some species may not be far behind^{8,9}. An era of doing real quantum science with ultracold molecules is now upon us. ■

Paul S. Julienne is at the Joint Quantum Institute, NIST and the University of Maryland, Gaithersburg, Maryland 20899–8423, USA.
e-mail: psj@umd.edu

1. Stuhl, B. K. *et al.* *Nature* **492**, 396–400 (2012).
2. Micheli, A., Brennen, G. K. & Zoller, P. *Nature Phys.* **2**, 341–347 (2006).
3. Baranov, M. A., Dalmonte, M., Pupillo, G. & Zoller, P. *Chem. Rev.* **112**, 5012–5061 (2012).
4. Ospelkaus, S. *et al.* *Science* **327**, 853–857 (2010).
5. Quémener, G. & Julienne, P. S. *Chem. Rev.* **112**, 4949–5011 (2012).
6. Ni, K. K. *et al.* *Science* **322**, 231–235 (2008).
7. van de Meerakker, S. Y. T., Bethlem, H. L., Vanhaecke, N. & Meijer, G. *Chem. Rev.* **112**, 4828–4878 (2012).
8. Shuman, E. S., Barry, J. F. & DeMille, D. *Nature* **467**, 820–823 (2010).
9. Hummon, M. T. *et al.* Preprint at <http://arxiv.org/abs/1209.4069> (2012).

EARTH SCIENCE

Go with the lows

Before Father Christmas sets off from the North Pole, he will want to know if his flight will be disrupted by polar lows — storms (pictured) that afflict subpolar seas. Unfortunately for him, the effects of polar lows are not usually included in climate and seasonal forecasting models. Writing in *Nature Geoscience*, Condrón and Renfrew report that they should be (A. Condrón and I. A. Renfrew *Nature Geosci.* <http://dx.doi.org/10.1038/ngeo1661>; 2012).

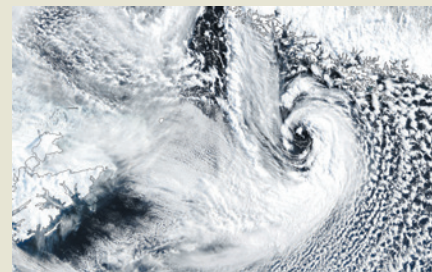
Polar lows are typically too small (less than 250 kilometres in diameter) and fleeting (24–48 hours in duration) to be well resolved in global meteorological and climate models. Nevertheless, their intensity is often high enough to affect convection in the underlying ocean. Deep open-ocean convection is one of the mechanisms that renews the North Atlantic Deep Water, the main water mass that drives large-scale

ocean circulation in the Atlantic Ocean.

Condrón and Renfrew used state-of-the-art computational models to simulate circulation in the northeast Atlantic during 1978–98, in the presence or absence of polar lows. They found that models that incorporated the effects of the storms indicated more open-ocean convection, at greater depths, than those that omitted polar lows.

Furthermore, the strengths of two gyres — large systems of rotating ocean currents — in the region were more frequently high when the effects of polar lows were simulated. This in turn increased the renewal of deep water in the Greenland Sea, the transport of that water flowing south across the Greenland–Iceland–Scotland ocean ridge, and the northward movement of heat to Europe and North America.

The authors conclude that the effects



of polar lows should be incorporated in ocean, climate and seasonal forecasting models. They also point out that the Intergovernmental Panel on Climate Change predicts that these storms will shift northwards in the future, and will occur less frequently than now. If so, this could greatly affect the deep waters of the North Atlantic, potentially reducing their southward flow. **Andrew Mitchinson**

2012

EDITORS' CHOICE

Adapted extracts from
selected News & Views
articles published this year.

CANCER

CLINICAL TRIALS UNITE MICE AND HUMANS

Leisa Johnson (*Nature* 483, 546–548; 2012)

In anticancer 'co-clinical' trials, mice carrying known mutations are treated in parallel with patients enrolled in a simultaneous clinical study. Chen and colleagues present a compelling case for conducting co-clinical trials. If done properly, co-clinical trials may help to identify predictive genetic markers that can be validated in real time using samples from patients enrolled in a concurrent clinical trial. These integrated data sets may ultimately be better at predicting the results of the concurrent clinical studies, as well as providing, on the basis of the cancer's genetic profile, a rationale for the observed differences in therapeutic response. Moreover, such coordinated processes could serve to inform the analysis and design of both current and future clinical trials, with a goal of increasing clinical success rates and decreasing health-care costs.

Nature 483, 613–617 (2012).

THERMODYNAMICS

THE FRIDGE GATE

Renato Renner (*Nature* 482, 164–165; 2012)

Minimalism is a popular trend in design, striving to expose the essence of an object through the elimination of all non-essential parts. Writing in the *Journal of Physics A*, Skrzypczyk *et al.* have now applied this approach to the study of thermal machines such as heat engines and refrigerators. Reducing the complexity of a refrigerator to its extreme, they arrived at a device as simple as a single logic gate. At first sight, refrigerators have little in common with such information-processing gates. But as early as the 1960s, information processing was studied from a thermodynamic perspective. It turns out that knowledge can always be traded for coldness. It is therefore no surprise that information-processing devices, such as logic gates, can have useful thermodynamic properties. The new minimalist fridge is a beautiful manifestation of this. Although it operates like a simple three-bit logic gate, it has the functionality of a fully fledged refrigerator.

J. Phys. A 44, 492002 (2011).

STRUCTURAL BIOLOGY

HOW OPIOID DRUGS BIND TO RECEPTORS

Marta Filizola & Lakshmi A. Devi (*Nature* 485, 314–317; 2012)

Opioid drugs such as morphine and codeine are powerful painkillers, but an assortment of adverse side effects limits their effective medical use. Four papers report crystal structures that provide the first direct



NASA, ESA & THE HUBBLE HERITAGE TEAM (STSC)

ASTRONOMY

COLLISION COURSE

R. Brent Tully (*Nature* 488, 600–601; 2012)

In a series of three papers published in *The Astrophysical Journal*, van der Marel and collaborators discuss the timing and dynamics of the 'imminent' — or at least inevitable — collision between the Milky Way and the neighbouring Andromeda galaxy. The question since 1959 has been whether the two galaxies would collide on first return or fly past each other. The first passage of Andromeda about the Milky Way is going to be close enough to make a big mess. Four billion years from now, our progeny will see, if they still have dark skies and keen eyes, quite a spectacle as Andromeda fills the horizon — just imagine being a resident in one of the members of the pair of colliding spiral galaxies NGC 2207 and IC 2163 (pictured).

Astrophys. J. 753, 7, 8, 9 (2012).

evidence for the binding mode of opioids to their receptors. The papers present the long-awaited, high-resolution crystal structures of all four opioid receptors in ligand-bound conformations. To develop drugs that retain the therapeutic action of opioids but not the unwanted side effects, it is crucial to understand the specific receptor conformations that opioids stabilize to selectively activate signalling pathways. This important aspect of ligand binding to opioid receptors is not captured by the crystal structures, and should be the subject of future research. Nevertheless, these crystal structures of inactive opioid receptors will contribute crucial information to a broad range of therapeutic areas, including those focused on pain, addiction and mental disorders.

Nature 485, 321–326, 327–332, 395–399, 400–404 (2012).

CLIMATE SCIENCE

AEROSOLS AND ATLANTIC ABERRATIONS

Amato Evan (*Nature* 484, 170–171; 2012)

Over the past century, the surface of the North Atlantic Ocean has gone through warm and cool periods that are not observed in other ocean basins. This Atlantic Multidecadal Oscillation (AMO) is thought to affect climate processes ranging from the current high levels of Atlantic hurricane activity to the devastating sub-Saharan droughts of the early 1980s. Booth *et al.* report their use of a state-of-the-art model of Earth's climate to demonstrate that, at least over the past century, the AMO is largely the response of the upper ocean to changes in the concentration of pollution aerosols in the atmosphere. The authors' evidence is compelling,

366 DAYS: the year in science

but their results are sensitive to model parameterizations of microphysical processes, particularly the interaction between cloud water droplets and aerosols, that are not well constrained by observations. If the results can be corroborated, then they suggest that multidecadal temperature fluctuations of the North Atlantic are dominated by human activity.

Nature **484**, 228–232 (2012).

BIODIVERSITY REMOTE RESPONSIBILITY

Edgar Hertwich (*Nature* **486**, 36–37; 2012)

If you buy a set of chess figures carved from ivory, you can suspect that you have contributed to killing an elephant. But if you buy a sausage, you cannot know whether the pig that was turned into the sausage was fed soy meal sourced from a farm that had just expanded into elephant habitat. The effects on species diversity, however, are similar. Understanding the complete causality chains leading to animal species extinctions has proven an intractable problem. Lenzen and colleagues present an analysis of species threats associated with international traded commodities, based on a detailed model of the global supply chains that connect final consumption to economic activities. Their results indicate that 30% of instances of red-listed species worldwide are caused by internationally traded commodities, and that the United States, Japan and European countries are the main net ‘importers’ of species threats, whereas south-east Asian countries are the main net ‘exporters’.

Nature **486**, 109–112 (2012).

NANOTECHNOLOGY THE IMPORTANCE OF BEING MODULAR

Paul W. K. Rothemund & Ebbe Sloth Andersen
(*Nature* **485**, 584–585; 2012)

Carpenters have been turning trees into furniture and dwellings for thousands of years, and so the discipline of woodworking features well-established techniques. Nanotechnologists similarly try to use DNA as a material for crafting nanometre-scale shapes, but DNA-working techniques are still evolving. Wei *et al.* present a method whose intrinsic modularity enables arbitrary DNA shapes to be constructed with striking speed. In their system, each tile is a single DNA strand with four



different binding domains that specify which four other tiles can bind to it as neighbours. The authors’ general scheme specifies a set of *N* tiles that self-assemble to form a rectangle, within which each tile adopts a particular position. By mixing together appropriate subsets of tiles and allowing them to self-assemble, arbitrary DNA shapes (pictured) can be prepared.

Careful studies of yields, kinetics and mechanism will be required to circumscribe the conditions under which the method works best.

Nature **485**, 623–626 (2012).

MICROBIOLOGY FAT, BILE AND GUT MICROBES

Peter J. Turnbaugh (*Nature* **487**, 47–48; 2012)

Humans never truly dine alone — our diet is intimately linked to the functioning of the trillions of microbes that inhabit our gastrointestinal

tract, the gut microbiota. Devkota and colleagues’ work emphasizes the importance of viewing nutrition from a perspective that encompasses both our human and microbial genomes. As the authors elegantly show, diets that provide the same number of calories can have remarkably different effects depending on the type of fat. It will be interesting to identify the specific components of our diet that influence microbial community structure, and to find out whether diets rich in saturated fat can drive the expansion of *Bilophila wadsworthia* or other potentially harmful microbes in humans. In this study, diet-driven changes in the production of bile acids affected gut microbes that, in turn, triggered disease. This line of research could ultimately lead to dietary recommendations tailored to match the idiosyncrasies of each person’s gut microbiota.

Nature **487**, 104–108 (2012).

FORUM Agriculture

COMPARING APPLES WITH ORANGES

(*Nature* **485**, 176–177; 2012)

A meta-analysis of agricultural systems shows that organic yields are mostly lower than those from conventional farming, but that organic crops perform well in some contexts.



THE FRUITS OF ORGANIC FARMING

John P. Reganold

Seufert and colleagues’ study is an example of meta-analysis being a great tool for identifying broad patterns not immediately visible in primary field research. If we want to feed a growing world population, producing adequate crop yields is vital, and this study bolsters the argument that adoption of organic agriculture under conditions in which it performs best might close the yield gap between organic and conventional systems. Organic farming provides multiple sustainability benefits, and these findings indicate that it can play a part in feeding the world.

GETTING BACK TO THE FIELD

Achim Dobermann

It is time to accept that various types of agriculture can have a place in feeding the world, but we also need to leave vague, outdated concepts of sustainability behind. Organic or low-external-input agriculture is not always sustainable. There are also many conventional agricultural systems that are highly productive, resource-efficient and sustainable. Instead of doing further meta-analyses to attempt to determine the optimal combination of agricultural systems, scientists should return to their fields and laboratories, and concentrate their efforts on increasing the performance of both conventional and organic agriculture.

Nature **485**, 229–232 (2012).

FOODCOLLECTION RF/GETTY IMAGES

Seventy-five genetic loci influencing the human red blood cell

A list of authors and their affiliations appears at the end of the paper

Anaemia is a chief determinant of global ill health, contributing to cognitive impairment, growth retardation and impaired physical capacity. To understand further the genetic factors influencing red blood cells, we carried out a genome-wide association study of haemoglobin concentration and related parameters in up to 135,367 individuals. Here we identify 75 independent genetic loci associated with one or more red blood cell phenotypes at $P < 10^{-8}$, which together explain 4–9% of the phenotypic variance per trait. Using expression quantitative trait loci and bioinformatic strategies, we identify 121 candidate genes enriched in functions relevant to red blood cell biology. The candidate genes are expressed preferentially in red blood cell precursors, and 43 have haematopoietic phenotypes in *Mus musculus* or *Drosophila melanogaster*. Through open-chromatin and coding-variant analyses we identify potential causal genetic variants at 41 loci. Our findings provide extensive new insights into genetic mechanisms and biological pathways controlling red blood cell formation and function.

Haemoglobin, an iron-containing metalloprotein found in the red blood cells of all vertebrates, provides the primary mechanism for oxygen transport in the circulation. Haemoglobin levels and related red blood cell phenotypes are tightly regulated, including an important genetic component^{1–5}. To refine our understanding of the genetic factors influencing red blood cell formation and function, we carried out a meta-analysis of genome-wide association studies (GWAS) and staged follow-up genotyping of six red blood cell phenotypes: haemoglobin, mean cell haemoglobin (MCH), mean cell haemoglobin concentration (MCHC), mean cell volume (MCV), packed cell volume (PCV) and red blood cell count (RBC).

Our study design is summarized in Supplementary Fig. 1. In brief, we combined genome-wide association data from 71,861 individuals of European or South Asian ancestry, with up to 2,644,161 autosomal single-nucleotide polymorphisms (SNPs) and 67,645 X-chromosome SNPs. Characteristics of participants, genotyping arrays and imputation are summarized in Supplementary Tables 1–3. Meta-analysis was carried out among Europeans and South Asians separately, followed by a final combined analysis of results from the two populations. We performed replication testing of 22 loci showing suggestive association ($10^{-8} < P < 10^{-7}$) in a further 63,506 individuals using a combination of *in silico* data and direct genotyping (Supplementary Tables 1, 2 and Supplementary Note). Genome-wide significance was set at $P < 1 \times 10^{-8}$, allowing a Bonferroni correction both for the $\sim 10^6$ independent SNPs tested⁶, as well as for the six inter-related red blood cell phenotypes (Supplementary Note)⁷.

Seventy-five independent genetic loci reached genome-wide significance for association with one or more red blood cell phenotypes (Table 1 and Supplementary Fig. 2), 43 of which are novel. For descriptive and downstream purposes, we identified a single ‘sentinel’ SNP for each of the 75 loci, defined as the SNP with the lowest P value against any phenotype at each locus; regional plots for the 75 loci are shown in Supplementary Fig. 3. Full lists of the SNPs associated with phenotype at $P < 10^{-6}$ and of the sentinel SNPs are provided (Supplementary Tables 4 and 5). Of the 38 loci previously reported to be associated with red blood cell traits^{1–5}, we replicate 32 loci ($P < 10^{-8}$) and find three to be nominally associated ($P < 0.05$; Supplementary Table 6). The remaining three loci, initially reported in an East Asian GWAS⁴, were not associated with red blood cell

phenotypes in our sample (Supplementary Fig. 4 and Supplementary Note).

Among the 75 genomic loci identified, we found that 31 were associated with one red blood cell phenotype, and 44 with two or more phenotypes, at $P < 10^{-8}$. The total number of locus–phenotype associations identified at $P < 10^{-8}$ was 156, of which 92 are novel (Supplementary Fig. 5 and Supplementary Table 7). In addition, at 8 of the 75 loci we found evidence for multiple SNPs independently associated with red blood cell phenotype at $P < 10^{-8}$ in conditional analyses⁸, suggesting the presence of possible secondary genetic mechanisms at these loci (Supplementary Table 8).

Identification of candidate genes

There are >3,000 protein-coding genes within 1 megabase (Mb) of the sentinel SNPs from the 75 genetic loci associated with red blood cell phenotypes. We prioritized genes as probable candidates underlying the observed genetic associations using the following criteria: (1) gene nearest to the sentinel SNP, and any other gene within 10 kilobases (kb) (97 genes; Table 1); (2) gene containing a non-synonymous SNP in high linkage disequilibrium ($r^2 > 0.8$) with the sentinel SNP (24 genes; Supplementary Table 9); (3) gene with expression quantitative trait loci (eQTL) associated with sentinel SNP in peripheral blood lymphocytes (27 genes; Supplementary Table 10); and (4) gene relationships among implicated loci (GRAIL) literature analysis⁹ (9 genes; Supplementary Table 11). This strategy identified 121 candidate genes (Table 1 and Supplementary Fig. 6).

Pathway analysis revealed that the list of candidate genes is strongly enriched for genes known to be involved in haematological development and function ($P = 10^{-63}$), as well as in cellular proliferation, development and death, and immunological processes (Supplementary Tables 12 and 13). Current knowledge of gene function for all 121 candidates is summarized in Supplementary Table 14. Of note, some of the genes within these regions are known to underlie the Mendelian red blood cell disorders of elliptocytosis, ovalocytosis and spherocytosis (*ANK1*, *SLC4A1*, *SPTA1*)¹⁰, haemolytic anaemia (*HK1*)¹¹ and iron deficiency or overload (*TMPRSS6*, *HFE*, *TFR2*)¹². Furthermore, somatic mutations of *IKZF1*, *KIT*, *SH2B3*, *SH3GL1* and *TAL1* (also known as *SCL*) underlie several haematologic proliferative disorders (Supplementary Table 14).

Table 1 | Genomic loci associated with red blood cell phenotypes

Region	Sentinel SNP	Position (B36)	Alleles (EA/OA)	EAF	Phenotype	Effect (SE)	P	Candidate genes
1p36	rs1175550	3,681,388	G/A	0.22	MCHC	0.008 (0.013)	8.6×10^{-15}	CCDC27 ⁿ , LRRC48 ⁿ
1p34†	rs3916164	39,842,526	G/A	0.71	MCH	0.008 (0.004)	3.1×10^{-10}	HEYL ⁿ
1p32	rs741959	47,448,820	G/A	0.57	MCV	0.157 (0.025)	6.0×10^{-10}	TAL1 ⁿ
1q23†	rs857684	156,842,353	C/T	0.74	MCHC	-0.006 (0.011)	3.5×10^{-16}	OR6Y1 ^c , OR10Z1 ^{nc} , SPTA1 ^{ncg}
1q32†	rs7529925	197,273,831	C/T	0.28	RBC	0.014 (0.002)	8.3×10^{-9}	MIR181A1 ⁿ
1q32	rs7551442	201,921,744	A/G	0.09	MCHC	-0.023 (0.017)	9.7×10^{-12}	ATP2B4 ^{ng}
1q32	rs9660992	203,516,073	G/A	0.42	MCH	0.007 (0.004)	7.1×10^{-10}	TMCC2 ⁿ
1q44†	rs3811444	246,106,074	T/C	0.35	RBC	0.018 (0.003)	4.5×10^{-10}	TRIM58 ^{nc}
2p21†	rs4953318	46,208,555	A/C	0.62	PCV	0.152 (0.018)	3.1×10^{-19}	PRKCE ⁿ
2p16†	rs243070	60,473,790	T/A	0.72	MCV	-0.181 (0.027)	4.4×10^{-13}	BCL11A ⁿ
2q13	rs10207392	111,566,130	G/A	0.44	MCV	-0.132 (0.025)	$4.4 \times 10^{-11*}$	ACOXL ⁿ
3p24†	rs9310736	24,325,815	A/G	0.35	MCV	-0.210 (0.026)	6.1×10^{-16}	THRB ⁿ
3q22	rs6776003	142,749,183	A/G	0.44	MCV	-0.138 (0.026)	$3.7 \times 10^{-11*}$	RASA2 ⁿ
3q23	rs13061823	143,603,476	T/C	0.56	MCV	-0.168 (0.025)	4.7×10^{-13}	XRNI ⁿ
3q29†	rs11717368	197,318,754	C/G	0.52	MCH	0.008 (0.004)	6.6×10^{-19}	TFRC ^{ng}
4q11†	rs218238	55,089,781	A/T	0.78	RBC	0.033 (0.003)	2.8×10^{-39}	KIT ⁿ
4q27	rs13152701	122,970,511	A/G	0.37	MCV	0.150 (0.026)	9.0×10^{-10}	BBS7 ⁿ , CCNA2 ^{ne}
6p23	rs6914805	16,389,166	C/T	0.75	MCH	0.012 (0.004)	1.2×10^{-19}	GMPPR ^{ne}
6p21†	rs1408272	25,950,930	G/T	0.07	MCH	0.033 (0.009)	4.8×10^{-67}	HFE ^{cg} , SLC17A3 ⁿ
6p22	rs13219787	27,969,649	A/G	0.09	MCH	0.023 (0.007)	5.9×10^{-17}	HIST1H2AM ⁿ , HIST1H2BO ⁿ , HIST1H3J ⁿ
6p22	rs2097775	30,462,282	A/T	0.15	HB	0.055 (0.008)	1.3×10^{-10}	TRIM39-RPP21 ⁿ
6p21	rs9272219	32,710,247	G/T	0.72	RBC	0.015 (0.002)	4.3×10^{-10}	HLA-DQA1 ^{nc} , HLA-DQA2 ^e
6p21†	rs9349204	42,022,356	G/A	0.27	MCV	-0.367 (0.028)	2.4×10^{-40}	CCND3 ⁿ
6p12	rs9369427	43,919,408	A/C	0.68	HB	0.042 (0.006)	5.6×10^{-12}	VEGFA ⁿ
6q21†	rs1008084	109,733,658	G/A	0.56	MCH	-0.010 (0.003)	6.4×10^{-26}	CCDC162P ⁿ
6q23†	rs9389269	135,468,852	T/C	0.72	MCV	-0.600 (0.028)	2.6×10^{-19}	HBS1L ⁿ
6q24†	rs590856	139,886,122	G/A	0.43	MCV	0.313 (0.026)	5.0×10^{-36}	CITED2 ⁿ
6q26	rs736661	164,402,826	A/G	0.62	MCH	0.007 (0.004)	1.6×10^{-11}	QKI ⁿ
7p13†	rs12718598	50,395,939	T/C	0.51	MCV	-0.204 (0.030)	1.6×10^{-13}	IKZF1 ⁿ
7q22†	rs2075672	100,078,232	A/G	0.39	RBC	0.022 (0.003)	1.9×10^{-20}	ACTL6B ⁿ , TFR2 ^{ng}
7q36†	rs10480300	151,036,938	C/T	0.72	HB	0.052 (0.007)	7.8×10^{-15}	PRKAG2 ^{ng}
8p11	rs4737009	41,749,562	G/A	0.74	MCHC	-0.014 (0.013)	4.9×10^{-11}	ANK1 ^{ng}
8p11	rs6987853	42,576,607	C/T	0.62	MCHC	-0.002 (0.010)	6.1×10^{-11}	C8orf40 ^{ne}
9p24†	rs2236496	4,834,265	C/T	0.22	MCV	-0.279 (0.031)	1.4×10^{-19}	RCL1 ⁿ
9q34†	rs579459	135,143,989	T/C	0.8	RBC	0.021 (0.003)	9.3×10^{-18}	ABO ⁿ
10q11†	rs901683	45,286,428	A/G	0.08	MCV	0.364 (0.050)	1.5×10^{-16}	MARCH8 ^{nce}
10q22†	rs10159477	70,769,894	A/G	0.16	HB	0.087 (0.010)	4.4×10^{-20}	HK1 ^{ng}
10q24	rs11190134	101,272,190	G/A	0.6	MCH	-0.011 (0.004)	$1.3 \times 10^{-10*}$	NKX2-3 ⁿ
11p15	rs11042125	8,894,625	A/T	0.6	HB	0.032 (0.006)	1.5×10^{-9}	AKIP1 ^{ne} , C11orf16 ^{ne} , NRIP3 ^e , ST5 ⁿ
11p15	rs7936461	9,997,462	C/T	0.75	PCV	0.121 (0.021)	1.0×10^{-9}	SBF2 ⁿ
11q13	rs2302264	66,964,002	G/A	0.58	MCV	0.140 (0.025)	1.3×10^{-10}	CORO1B ^{ne} , PTPRCAP ^{ne} , RPS6KB2 ^{nce}
11q13	rs7125949	72,686,732	A/G	0.11	HB	0.053 (0.010)	2.1×10^{-9}	ARHGEF17 ^{ce} , P2RY6 ⁿ
12p13	rs7312105	2,393,616	G/A	0.36	PCV	0.104 (0.019)	$3.2 \times 10^{-9*}$	CACNA1C ⁿ
12p13†	rs10849023	4,202,739	C/T	0.79	MCH	-0.008 (0.005)	7.5×10^{-12}	CCND2 ^{ng}
12q22	rs11104870	87,353,425	C/T	0.3	RBC	0.013 (0.002)	$6.2 \times 10^{-11*}$	KITLG ⁿ
12q24†	rs3184504	110,368,991	T/C	0.48	HB	0.051 (0.006)	4.3×10^{-19}	ATXN2 ⁿ , SH2B3 ^{nc}
12q24	rs3829290	119,610,821	C/T	0.44	MCV	-0.153 (0.026)	2.1×10^{-9}	ACADS ^c , MLEC ⁿ
14q23†	rs7155454	64,571,992	A/G	0.51	MCH	0.002 (0.004)	1.8×10^{-12}	FNTB ⁿ , MAX ⁿ
14q24	rs11627546	69,435,677	C/A	0.84	MCV	0.162 (0.032)	$1.1 \times 10^{-9*}$	SMOC1 ⁿ
14q32†	rs17616316	102,892,515	G/C	0.07	MCH	0.014 (0.009)	$8.2 \times 10^{-11*}$	EIF5 ⁿ
15q21†	rs1532085	56,470,658	G/A	0.59	HB	0.034 (0.006)	$6.7 \times 10^{-11*}$	LIPC ⁿ
15q22†	rs2572207	63,857,747	C/T	0.74	MCV	0.153 (0.029)	3.4×10^{-9}	DENND4A ⁿ , PTPLAD1 ^e
15q24	rs8028632	73,108,315	T/C	0.8	MCV	0.188 (0.032)	6.9×10^{-10}	PPCDC ⁿ , SCAMP5 ⁿ
15q24	rs11072566	74,081,026	A/G	0.48	HB	0.028 (0.006)	$3.0 \times 10^{-10*}$	NRG4 ⁿ
15q25	rs2867932	76,378,092	G/A	0.61	MCHC	-0.021 (0.010)	3.3×10^{-9}	DNAJA4 ^e , WDR61 ⁿ
16p11†	rs11248850	103,598	G/A	0.5	MCH	0.007 (0.004)	6.3×10^{-23}	NRLP3 ⁿ
16q22	rs2271294	66,459,827	T/A	0.15	RBC	0.017 (0.003)	1.1×10^{-9}	CTRL ^c , DUS2L ^e , EDC4 ⁿ , NUTF2 ⁿ , PSMB10 ^c
16q24†	rs10445033	87,367,963	G/A	0.37	MCHC	0.020 (0.012)	1.5×10^{-22}	PIEZO1 ⁿ
17p11	rs888424	19,926,019	A/G	0.43	MCH	0.006 (0.004)	5.4×10^{-20}	SPECC1 ⁿ
17q11	rs2070265	24,099,550	T/C	0.2	MCH	0.013 (0.004)	5.1×10^{-14}	C17orf63 ⁿ , ERAL1 ^e , NEK8 ⁿ , TRAF4 ^{ne}
17q12	rs8182252	34,981,476	C/T	0.18	RBC	0.016 (0.003)	5.9×10^{-9}	CDK12 ^e , NEUROD2 ⁿ
17q21	rs2269906	39,649,863	C/A	0.36	MCHC	0.027 (0.010)	2.0×10^{-11}	SLC4A1 ^e , UBTFT ⁿ
17q21	rs12150672	41,182,408	A/G	0.23	RBC	0.017 (0.003)	4.7×10^{-12}	ARHGAP27 ^e , ARL17B ^e , C17orf69 ^{ce} , CRHR1 ^{nc} , SPPL2C ^c , KANSL1 ^c , MAPT ^c , STH ^c
17q25	rs4969184	73,905,008	G/A	0.53	HB	0.031 (0.006)	7.0×10^{-9}	PGS1 ^{ne}
18q21	rs4890633	42,087,276	G/A	0.27	MCH	0.005 (0.004)	1.9×10^{-23}	C18orf25 ^{ne}
19p13	rs2159213	2,087,102	C/T	0.5	HB	0.032 (0.006)	1.9×10^{-9}	AP3D1 ⁿ
19p13	rs732716	4,317,219	A/G	0.71	MCV	0.201 (0.028)	1.5×10^{-14}	MPND ⁿ , SH3GL1 ⁿ , UBXLN6 ^c
19p13†	rs741702	12,885,250	A/C	0.35	MCH	0.006 (0.004)	8.2×10^{-20}	CALR ^e , FARSA ^{ne} , SYCE2 ⁿ
19q13	rs3892630	37,873,324	T/C	0.18	MCV	0.176 (0.034)	$1.0 \times 10^{-10*}$	NUDT19 ^{nc}
20q13†	rs737092	55,423,811	C/T	0.49	MCV	0.216 (0.033)	4.0×10^{-13}	RBMS38 ⁿ
21q22†	rs2032314	34,276,393	T/C	0.08	PCV	0.154 (0.034)	$7.5 \times 10^{-10*}$	ATP5O ⁿ
22q11†	rs5754217	20,269,675	G/T	0.83	MCV	0.194 (0.031)	8.6×10^{-10}	UBE2L3 ^{ne} , YDJC ^c
22q12†	rs5749446	31,210,585	T/C	0.62	MCH	0.007 (0.004)	3.3×10^{-13}	FBXO7 ^{ncg}
22q12†	rs855791	35,792,882	G/A	0.57	MCH	0.012 (0.004)	1.0×10^{-69}	KCTD17 ⁿ , TMPPRSS6 ^{nc}
22q13†	rs140522	49,318,132	C/T	0.67	MCV	0.287 (0.030)	4.5×10^{-23}	TYMP ^{ne} , NCAPH2 ⁿ , ODF3B ⁿ , SCO2 ⁿ

Candidate gene superscripts indicate the method of identification. *Replication testing performed. †Previously reported. ‡Discovered from combined analysis of European and South Asian genome-wide association data. c, coding variant; e, eQTL; EA, effect allele; EAF, effect allele frequency; g, GRail; HB, haemoglobin; n, nearby; OA, other allele; SE, standard error.

Gene expression during haematopoiesis

We next explored expression of the 121 candidate genes using an atlas of 38 different haematopoietic cell types (Supplementary Table 15)¹³. Ninety-seven genes could be reliably assigned a probe on the Affymetrix HG_U133AAofAv2 array (Fig. 1a); these transcripts were, on average, expressed at higher levels in late erythroblasts (or the precursors of red blood cells, EB3-EB5) compared to other transcripts in the same cell type ($P < 0.01$ after Bonferroni correction; Fig. 1b). Furthermore, expression was more likely to be upregulated in EB3-5 relative to other cell types ($P = 1.2 \times 10^{-6}$, rank-sum test).

To further investigate lineage-specific effects, we assessed temporal patterns of gene expression during *in vitro* differentiation of haematopoietic stem cells to erythroblasts¹⁴. On average, candidate genes have increasing expression over time along the erythroid lineage ($P = 0.006$, rank-sum test; Fig. 1c). These data support the view that the gene set identified here is enriched for genes relevant to red blood cell biology, including a number of candidate genes differentially regulated to increase their expression in late erythropoiesis.

Coding and regulatory sequence variants

To better capture common sequence variation at the 75 loci, we searched the 1000 Genomes Project data set (www.1000genomes.org)

and identified 39 non-synonymous SNPs that are in high linkage disequilibrium ($r^2 > 0.8$) with sentinel SNPs at the red blood cell loci (Supplementary Table 9). This represents a \sim sixfold enrichment compared to the expectation under the null hypothesis ($P = 0.01$; Supplementary Note). Although re-sequencing will be needed to obtain a complete assessment of variants at these loci, these non-synonymous sites represent an initial set of candidates for genetic variants underlying the observed associations with red blood cell phenotypes, potentially mediated through changes in protein function.

We next searched for sequence variants at the red blood cell loci that might influence gene regulation. We used formaldehyde-assisted isolation of regulatory elements followed by next-generation sequencing (FAIRE-seq) to identify nucleosome-depleted regions (NDRs) that may represent active regulatory elements¹⁵. We studied three haematologic cell types, and found 103,308 unique NDRs, of which 38,014 were present in erythroblasts, 50,372 in megakaryocytes and 34,833 in monocytes. We then searched the 1000 Genomes Project data set and found 60 SNPs located within one of these NDRs that are either: (1) one of the 75 sentinel SNPs from the red blood cell GWAS, or (2) in high linkage disequilibrium ($r^2 > 0.8$) and located within 1 Mb of a sentinel SNP (Supplementary Table 16). The NDRs overlapping these 60 SNPs were more likely to be erythroblast specific than expected by

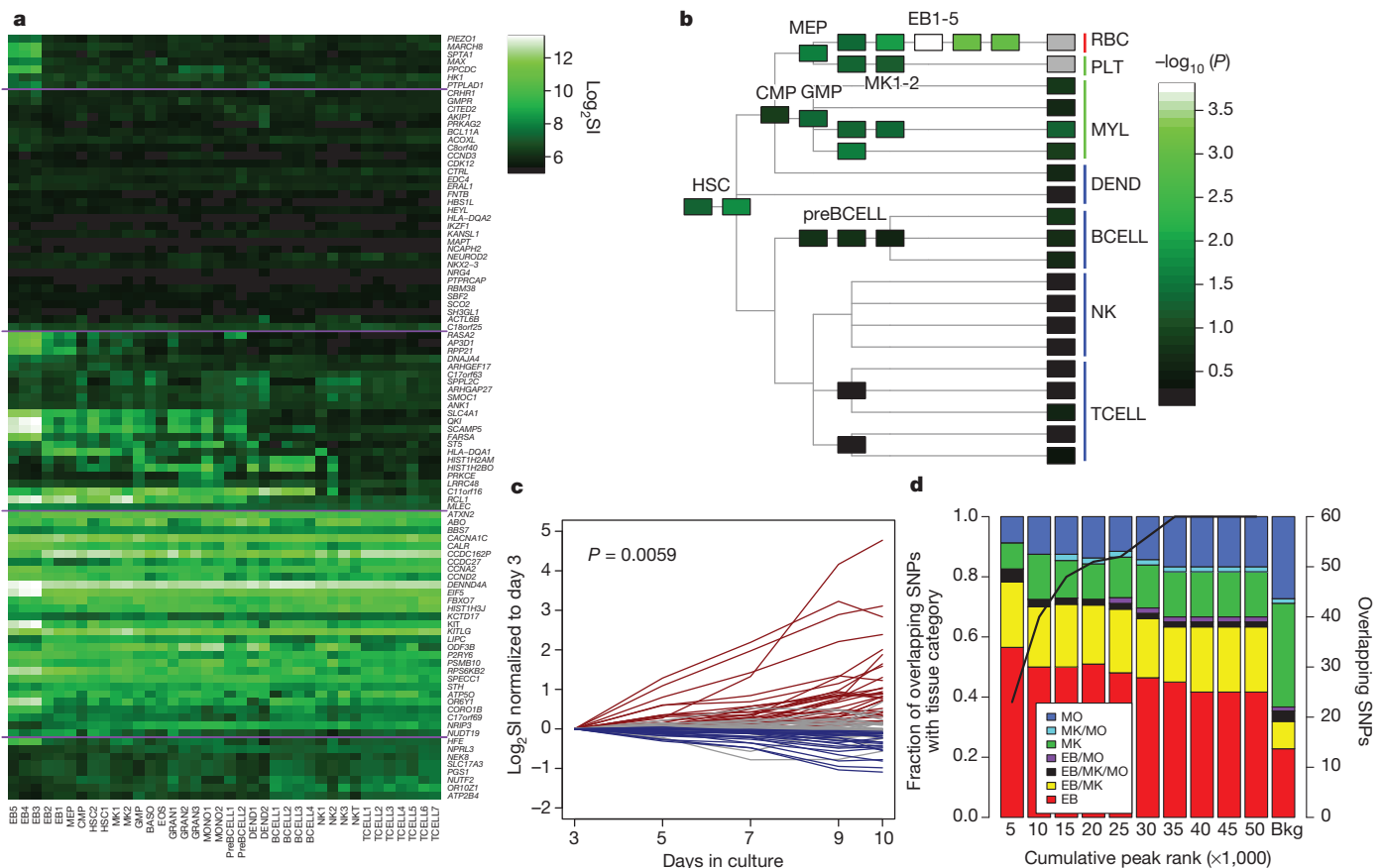


Figure 1 | Gene-expression patterns for 121 putative candidate genes, and tissue distribution of NDRs. **a**, Heat-map of candidate genes in the Differentiation Map of Hematology¹³. Cell acronyms refer to original source (summarized in Supplementary Table 15). Expression above a \log_2 signal intensity (SI) of 6 is consistently above background. **b**, $-\log_{10} P$ of the signed-rank test for candidate genes being more highly expressed in each cell type than non-candidate genes. **c**, Time-course of differentiation of cord-blood haematopoietic stem cells cultured along the erythroid lineage. Putative candidate genes are shown as upregulated (red), downregulated (blue) or with the slope not being significantly different from zero (grey). **d**, Tissue distribution of NDRs containing a potential causal variant. NDRs were ranked by peak score (proportional to their peak height in FAIRE-seq). The rankings were then used

to divide the NDRs into cumulative tranches to explore the effect of calling-thresholds on results (left bar, tranche containing the 5,000 top-ranked NDRs of each cell type; penultimate bar, tranche containing the 50,000 top-ranked NDRs of each cell type). The solid line indicates the number of SNPs overlapping the tranche-specific NDRs that are potential causal variants (defined as a sentinel SNP from the red blood cell GWAS, or a SNP in high linkage disequilibrium ($r^2 > 0.8$) and located within 1 Mb of a sentinel SNP; right-hand y axis); the bar summarizes the tissue distribution of these SNPs (as a percentage of tranche-specific total). The right-hand bar represents the expected tissue distribution for the SNPs under the null hypothesis. Results show that the potential causal variants are most commonly found in erythroblast-specific NDRs, and that this is true across the spectrum of peak-calling thresholds.

chance (1.8-fold enrichment compared to background distribution of NDRs; $P = 0.007$, Bonferroni-adjusted binomial test); by contrast, there were fewer megakaryocyte-specific NDRs coinciding with red blood cell SNPs (0.4-fold enrichment; $P = 0.007$; Fig. 1d). This pattern of erythroblast enrichment and megakaryocyte depletion was robust to the stringency of NDR peak-calling (Supplementary Table 17). Our results indicate that regulatory variation within the erythroid lineage may underlie the associations observed at several of the loci identified in our red blood cell GWAS. The 19 genes closest to the 25 erythroblast-specific NDRs were more likely to be upregulated during erythropoiesis compared to all other expressed transcripts ($P = 6.3 \times 10^{-6}$, rank-sum test; Supplementary Table 18), lending further support to the view that the NDRs identified have a role in the regulation of genes involved in erythropoiesis^{16,17}. Interestingly, the SNPs associated with MCH at 16p11 overlap an erythroblast-specific NDR that coincides with the NPRL3 regulatory element in the locus control region of the downstream haemoglobin- α locus^{18,19}.

Together our coding- and regulatory-variant analyses thus identify a set of ~100 SNPs across 41 regions that are candidates for functional genomic elements influencing red blood cell formation and function, and which constitutes a priority set for future experimental evaluation.

Insights from mouse models

A systematic search of the Mouse Genome Informatics database reveals haematologic phenotypes for 29 of the 100 candidate genes that have mouse homologues (Supplementary Fig. 6 and Supplementary Tables 14, 19), including genes involved in cell cycle regulation: *CCNA2* (4q27), *CCND2* (12p13) and *CCND3* (6p21); genes coding for transcription factors and their interacting proteins: *BCL11A* (2p16), *CITED2* (6q24), *IKZF1* (7p13) and *TAL1* (1p32); and genes involved in growth factor or cytokine signalling: *KIT* (4q11), *KITLG* (12q22), *SH2B3* (12q24) and *PTPRCAP* (11q13). Among the gene products encoded at the newly identified loci, *KITLG*, also known as stem cell factor, is the cognate ligand for the *KIT* tyrosine kinase receptor²⁰. *KIT* signalling is involved in the perinatal transition from fetal to adult haemoglobin, in addition to maintenance, proliferation and differentiation of haematopoietic stem cells²¹. *Kitlg*^{-/-} and *Kit*^{-/-} mice have low red blood cell concentrations, anaemia and other haematological abnormalities. *CCNA2*, *CCND2* and *CCND3* are cyclin-dependent kinases that contribute to initiation and progression of cell division²². Knock-out models of these genes have a number of haematological abnormalities, including reduced stem cell and red blood cell concentrations, and anaemia²². Of the 29 candidate genes with a blood phenotype in mouse, 25 were identified as the genes nearest to the sentinel SNP, and 15 through the eQTL ($n = 2$), coding-variant ($n = 6$) or GRAIL ($n = 8$) analyses (Supplementary Table 19).

RNAi silencing in *D. melanogaster*

We used haemocyte-specific RNA interference (RNAi) silencing in *D. melanogaster* to further evaluate the candidate genes for their role in blood cell formation. We first carried out permutation testing in a genome-wide *D. melanogaster* RNAi silencer screen (Supplementary Note). Results confirmed that the 121 candidates are enriched for genes with a blood cell phenotype in *D. melanogaster*, supporting the view that our GWAS identifies a set of genes conserved across phyla and involved in blood cell formation or survival.

We next created haemocyte-specific RNAi knockdowns for 96 *D. melanogaster* genes that are orthologues for 74 of the 121 candidate genes, and assessed blood cell formation (crystal cells and plasmotocytes) in early- and late-stage L3 larvae²³. We found 19 out of the 74 candidate genes with orthologues in *D. melanogaster* to have a blood cell phenotype, of which 5 also have a haematological phenotypes in mouse models: *KIT*, *HK1*, *CCNA2*, *AP3D1* and *PSMB10* (Supplementary Tables 19 and 20). Among the genes highlighted, RNAi silencing of *KIT* and *CCNA2* orthologues was associated with a profound reduction in plasmocyte formation (Fig. 2), consistent with

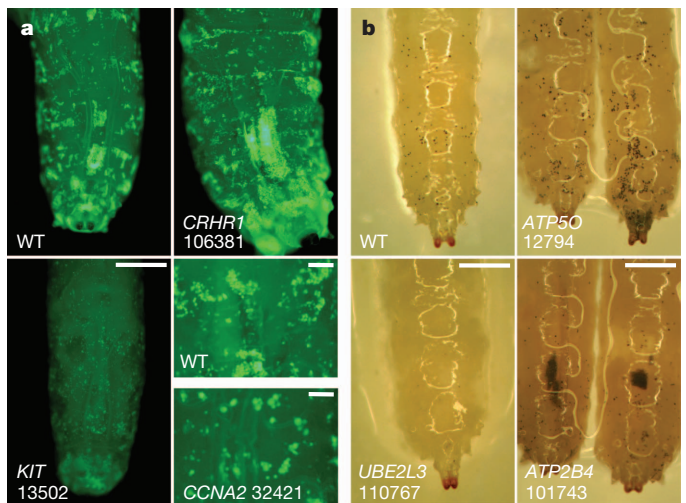


Figure 2 | RNAi silencing in *D. melanogaster*. **a**, Plasmotocytes imaged by green fluorescent protein expression (light green spots on posterior dorsal end of L3 larvae) from wild-type (WT) cells and cells with RNAi silencing of orthologues of the following human genes: *CRHR1* (106381, increased cell counts (CC)), *KIT* (13502, decreased CC) and *CCNA2* (32421, increased CC). Numbers represent the unique Flybase IDs corresponding to the *D. melanogaster* orthologues. Scale bar, 0.5 mm. Bottom right, plasmotocyte size is also increased in *CCNA2* compared to wild type. Scale bars, 0.1 mm. **b**, Crystal cells (black spots visualized by heating larvae to 60 °C) in wild-type larvae, and in RNAi silencing of *ATP5O* (12794, increased CC), *UBE2L3* (110767, decreased CC) or *ATP2B4* (101743, aggregated). Scale bars, 0.5 mm.

their established role in cytokinesis^{20,22}. *AP3D1* is involved in vesicular trafficking and dense granule formation in platelets²⁴, whereas *PSMB10* is a component of a widely distributed proteasome linked to inflammation and ubiquitin signalling²⁵. *UBE2L3* is also involved in ubiquitin signalling and immune regulation²⁶, and genetic variants in *UBE2L3* are strongly associated with several autoimmune diseases known to influence blood cell counts^{27,28}. *EIF5* (14q32) is involved in activation of the ribosomal initiation complex²⁹, whereas *RPS6KB2* (22q11) is a key component of growth factor and other signalling cascades that regulate ribosomal function, cellular proliferation and survival³⁰. For most of the genes identified, the mechanisms underlying their potential relationship to red cell biology remain to be elucidated; our gene set thus provides a rich resource for future experimental evaluation and discovery.

Contribution to clinical phenotype

The 75 sentinel SNPs together account for between 3.9% (PCV) and 8.9% (MCV) of population variation in red blood cell phenotypes (Supplementary Table 21). Individuals in the highest quartile of genetic risk score (GRS; on the basis of weighted effect of the 75 sentinel SNPs) are 3–5-fold more likely to be in the highest quartile for population distribution of MCH, MCV and RBC (Fig. 3). GRS is associated with haemoglobin concentrations across the physiological range, including at haemoglobin levels that predict adverse outcomes in pregnancy, cardiovascular and neurologic disease, in addition to mortality in the elderly^{31–34}.

We next investigated the association of the 75 sentinel SNPs with red blood cell phenotypes in thalassaemia, a group of genetic disorders characterized by defects in haemoglobin synthesis and anaemia. We confirmed association of several of the sentinel SNPs with respective blood cell trait, and found that GRS predicts phenotype similarly, among 460 β -thalassaemia heterozygotes (Supplementary Table 22 and Supplementary Note). In separate experiments, GRS predicts time to first blood transfusion among 495 patients with thalassaemia major ($P = 6.9 \times 10^{-4}$); however, this effect was fully accounted for by the *MYB-HBS1L* locus, which modifies the severity of thalassaemia major

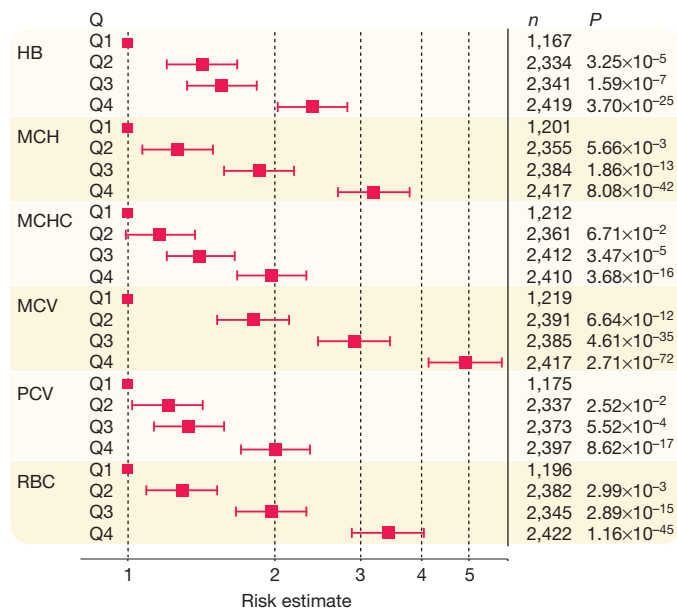


Figure 3 | Association of SNP score with red blood cell phenotypes. Results presented as odds ratio (95% confidence interval) for participants in each SNP score quartile (Q) having phenotype level in the top quartile versus the lowest quartile of the respective population distribution, compared to people in the lowest quartile of SNP score (Q1, reference group). HB, haemoglobin; n, number of participants in the respective comparison of SNP score quartiles.

through its effect on fetal haemoglobin levels (Supplementary Note)³⁵. Together, our findings demonstrate that the common genetic variants identified contribute to phenotypic variation in the general population, and suggest that they may also act as genetic modifiers in clinically relevant red blood cell abnormalities.

Conclusions

Our genome-wide association and replication study in 135,367 individuals identifies 75 genetic loci influencing red blood cell phenotypes, and 156 locus–phenotype associations; most of these discoveries are novel. Through open-chromatin and coding-variant studies, we identify a first set of SNPs as potential causal variants. In parallel, our bioinformatic strategies identify a core set of genes, differentially regulated in haematologic precursor cells, which are candidates for mediating the effects on red blood cell phenotypes. However, despite our extensive GWAS, bioinformatic and experimental data, the precise identities of the causal variants, regulatory regions and genes remain to be determined; definitive identification will require further detailed experimental evaluation. Our results thus provide new insights into the genes and gene variants that may influence haemoglobin levels and related red blood cell indices, and will underpin a deeper knowledge of the biological mechanisms involved in haematopoiesis and red blood cell function.

METHODS SUMMARY

Genome-wide association and replication. Genome-wide association was carried out in 62,553 people of European ancestry and 9,308 people of South Asian ancestry. Phenotypic associations were tested in each cohort separately, followed by fixed-effect meta-analysis using Z-scores weighted by the square root of sample size. Replication testing of 22 SNPs was done by *in silico* and direct genotyping among 63,506 people, and results combined with genome-wide association data. Genome-wide significance was inferred at $P < 1 \times 10^{-8}$.

Gene-expression profiling. Gene expression was investigated in cord-blood-derived CD34⁺ haematopoietic stem cells *in vitro*, differentiated along the erythroid lineage for 3, 5, 7, 9 or 10 days. Gene expression was assayed using Illumina human WGv3.0 microarrays, and temporal patterns quantified by linear regression.

Open-chromatin studies. FAIRE-seq was done in erythroblasts, megakaryocytes and peripheral blood monocytes. NDRs were identified as regions of sequencing

enrichment using F-Seq³⁶. Candidate functional SNPs were selected as all biallelic SNPs within 1 Mb of the sentinel SNP and in linkage disequilibrium at $r^2 > 0.8$. **D. melanogaster studies.** *D. melanogaster* orthologues of the human candidate genes were identified using Ensembl Compara. Haemocyte-specific RNAi silencing of the orthologues identified was achieved using the blood-specific hemolectin promoter driving the yeast transcriptional activator Gal4 (*Hml-Gal4*) which, in turn, promotes upstream activation sequence–short hairpin RNA expression. Early and late L3 larvae were analysed for plasmotocyte and crystal cell numbers and morphology.

Contribution to population variation. Phenotypic contribution was investigated in non-discovery samples. Estimates of population variance explained by the sentinel SNPs were made in each study separately, and mean values calculated weighted by sample size. We calculated the odds ratio for being in the highest versus the lowest quartile of phenotype, associated with a SNP score defined as the sum of number of effect (trait raising) alleles present, weighted according to effect size.

Full Methods and any associated references are available in the online version of the paper.

Received 6 February; accepted 15 October 2012.

Published online 5 December; corrected online 19 December 2012 (see full-text HTML version for details).

- Chambers, J. C. *et al.* Genome-wide association study identifies variants in *TM6SF2* associated with hemoglobin levels. *Nature Genet.* **41**, 1170–1172 (2009).
- Ganesh, S. K. *et al.* Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. *Nature Genet.* **41**, 1191–1198 (2009).
- Soranzo, N. *et al.* A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. *Nature Genet.* **41**, 1182–1190 (2009).
- Kamatani, Y. *et al.* Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nature Genet.* **42**, 210–215 (2010).
- Ding, K. *et al.* Genetic loci implicated in erythroid differentiation and cell cycle regulation are associated with red blood cell traits. *Mayo Clin. Proc.* **87**, 461–474 (2012).
- Pe'er, I., Yelensky, R., Altshuler, D. & Daly, M. J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385 (2008).
- Nyholt, D. R. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* **74**, 765–769 (2004).
- Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genet.* **44**, 369–375 (2012).
- Raychaudhuri, S. *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.* **5**, e1000534 (2009).
- An, X. & Mohandas, N. Disorders of red cell membrane. *Br. J. Haematol.* **141**, 367–375 (2008).
- van Wijk, R., Rijksen, G., Huizinga, E. G., Nieuwenhuis, H. K. & van Solinge, W. W. HK Utrecht: missense mutation in the active site of human hexokinase associated with hexokinase deficiency and severe nonspherocytic hemolytic anemia. *Blood* **101**, 345–347 (2003).
- Camaschella, C. & Poggiali, E. Inherited disorders of iron metabolism. *Curr. Opin. Pediatr.* **23**, 14–20 (2011).
- Novershtern, N. *et al.* Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296–309 (2011).
- Gieger, C. *et al.* New gene functions in megakaryopoiesis and platelet formation. *Nature* **480**, 201–208 (2011).
- Paul, D. S. *et al.* Maps of open chromatin guide the functional follow-up of genome-wide association signals: application to hematological traits. *PLoS Genet.* **7**, e1002139 (2011).
- Forrester, W. C., Thompson, C., Elder, J. T. & Groudine, M. A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc. Natl Acad. Sci. USA* **83**, 1359–1363 (1986).
- Tuan, D., Solomon, W., Li, Q. & London, I. M. The “beta-like-globin” gene domain in human erythroid cells. *Proc. Natl Acad. Sci. USA* **82**, 6384–6388 (1985).
- Kowalczyk, M. S. *et al.* Intragenic enhancers act as alternative promoters. *Mol. Cell* **45**, 447–458 (2012).
- Baù, D. *et al.* The three-dimensional folding of the α -globin gene domain reveals formation of chromatin globules. *Nature Struct. Mol. Biol.* **18**, 107–114 (2011).
- Zsebo, K. M. *et al.* Stem cell factor is encoded at the *Sl* locus of the mouse and is the ligand for the *c-kit* tyrosine kinase receptor. *Cell* **63**, 213–224 (1990).
- Heissig, B. *et al.* Recruitment of stem and progenitor cells from the bone marrow niche requires MMP-9 mediated release of kit-ligand. *Cell* **109**, 625–637 (2002).
- Kozar, K. *et al.* Mouse development and cell proliferation in the absence of D-cyclins. *Cell* **118**, 477–491 (2004).
- Dietzl, G. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151–156 (2007).
- Clark, R. H. *et al.* Adaptor protein 3-dependent microtubule-mediated movement of lytic granules to the immunological synapse. *Nature Immunol.* **4**, 1111–1120 (2003).

25. Berhane, S. *et al.* Adenovirus E1A interacts directly with, and regulates the level of expression of, the immunoproteasome component MECL1. *Virology* **421**, 149–158 (2011).
26. Tiwari, S. & Weissman, A. M. Endoplasmic reticulum (ER)-associated degradation of T cell receptor subunits. Involvement of ER-associated ubiquitin-conjugating enzymes (E2s). *J. Biol. Chem.* **276**, 16193–16200 (2001).
27. Franses, K. *et al.* Analysis of SNPs with an effect on gene expression identifies UBE2L3 and BCL3 as potential new risk genes for Crohn's disease. *Hum. Mol. Genet.* **19**, 3482–3488 (2010).
28. Zernakova, A. *et al.* Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet.* **7**, e1002004 (2011).
29. Das, S., Ghosh, R. & Maitra, U. Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. *J. Biol. Chem.* **276**, 6720–6726 (2001).
30. Fenton, T. R. & Gout, I. T. Functions and regulation of the 70 kDa ribosomal S6 kinases. *Int. J. Biochem. Cell Biol.* **43**, 47–59 (2011).
31. Scanlon, K. S., Yip, R., Schieve, L. A. & Cogswell, M. E. High and low hemoglobin levels during pregnancy: differential risks for preterm birth and small for gestational age. *Obstet. Gynecol.* **96**, 741–748 (2000).
32. Shah, R. C., Buchman, A. S., Wilson, R. S., Leurgans, S. E. & Bennett, D. A. Hemoglobin level in older persons and incident Alzheimer disease: prospective cohort analysis. *Neurology* **77**, 219–226 (2011).
33. Sabatine, M. S. *et al.* Association of hemoglobin levels with clinical outcomes in acute coronary syndromes. *Circulation* **111**, 2042–2049 (2005).
34. Zakai, N. A. *et al.* A prospective study of anemia status, hemoglobin concentration, and mortality in an elderly cohort: the Cardiovascular Health Study. *Arch. Intern. Med.* **165**, 2214–2220 (2005).
35. Galanello, R. *et al.* Amelioration of Sardinian β^0 thalassemia by genetic modifiers. *Blood* **114**, 3935–3937 (2009).
36. Boyle, A. P., Guinney, J., Crawford, G. E. & Furey, T. S. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* **24**, 2537–2538 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements A detailed list of acknowledgements is provided in the Supplementary Material.

Author Contributions Study organisation: J.C.C., C.G., P.v.d.H., J.S.K., W.H.O. and N.S. Manuscript preparation: H.A., J.S.B., J.C.C., G.V.D., P.D., C.G., P.v.d.H., A.A.Hicks, J.S.K., I.M.-L., W.H.O., A.Radhakrishnan, A.Rendon, S.S., J.Sehmi, N.S., D.S.P., M.U., N.V. and W.Z. All authors reviewed and had the opportunity to comment on the manuscript. Data collection and analysis in the participating genome-wide association, replication and phenotype cohorts: **ALSPAC**: D.M.E., J.P.K., S.M.R., G.D.S.; **AMISH**: Q.D.G., B.D.M., A.Parsa, A.R.S.; **Beta-thalassaemia**: F.A., F.D., P.Fortina, R.G., L.Perseu, A.Piga, S.S., M.U.; **CBR**: A.Attwood, J.D., S.F.G., H.L.-J., C.Moore, W.H.O., J.Sambrook; **CoLAUS**: F.B., J.S.B., M.H., P.V.; **DeCODE**: G.I.E., D.F.G., H.H., I.O., P.T.O., K.S., P.S., U.T.; **DESIR**: B.Balkau, C.D., P.Froguel, R.Sladek; **EGCUT**: T.E., K.F., A.M., E.M., A.S.; **EPIC**: K.-T.K., C.L., R.J.F.L., N.J.W., J.-H.Z.; **Genebank**: H.A., J.H., S.L.H., W.H.W.T.; **INGI CARL**: P.G., G.G., N.P.; **INGI CILENTO**: M.C., T.N., D.R., R.Sorice; **INGI FVG**: A.P.d.A., A.Robino, S.U.; **INGI Val Borbera**: G.P., C.S., D.T., M.T.; **KORA**: A.D., C.G., T.I., C.Meisinger, J.S.R.; **LBC**: I.J.D., S.E.H., L.M.L., J.M.S.; **LIFELINES**: R.A.d.B., I.P.K., I.M.-L., G.N., P.v.d.H., L.J.v.P., N.V., B.H.R.W.; **LLOIPOP**: A.A.Hussani, J.C.C., D.D., P.E., J.S.K., X.L., K.M., J.Scott, J.Sehmi, S.-T.T., W.Z.; **LURIC**: B.G., B.O.B., M.E.K., W.M., B.R.W.; **MDC**: A.F.D., G.E., B.H., C.E.H., O.M., S.P., J.G.S.; **MICROS**: M.G., A.A.Hicks, A.S.-P., P.P.P.; **NESDA**: I.M.N., B.W.P., J.H.S., H.Snieider; **NFBC1966**: A.-L.H., M.-R.J., P.F.O., A.Pouta, A.Ruokonen.; **NTR**: A.Abdellaoui, D.I.B., E.J.C.d.G., J.-J.H., M.H.d.M., G.Willemsen; **OGP**: F.M., D.P., L.Portas, M.P.; **PREVEND**: R.A.d.B., I.M.-L., G.N., P.v.d.H., W.H.v.G., D.J.v.V., N.V.; **QIMR**: B.Benjamin, M.A.F., N.G.M., S.E.M., G.W.M., C.S.T., P.M.V., J.B.W.; **Sardinia**: F.C., E.P., S.S., M.U.; **SHIP**: A.M., N.Auck, C.O.S., A.Teumer, U.V.; **SMART**: A.Algra, F.W.A., P.I.W.d.B., V.T.; **SORBS**: V.L., I.P., M.S., A.Tönjes; **TwinsUK**: Y.M., S.-Y.S., N.S., T.D.S.; **UKBS**: J.J., W.H.O., N.S., J.Stephens; **Young Finns**: M.K., T.L., L.-P.L., O.R. Functional studies: *Drosophila*, U.E., F.S.D., A.A.Hicks, M.Novatchkova, J.M.P., U.P., C.X.W., G.Wirnsberger; expression profiling, W.O.C., L.Franke, L.L., M.F.M., A.Rendon, E.S., H.-J.W.; **FAIRE**, C.A.A., P.D., W.H.O., D.S.P., A.Rendon, N.S. Data analysis and bioinformatics: A.A.Hussani, S.B., J.C.C., M.D., L.Ferrucci, P.v.d.H., S.K., X.L., I.M.-L., K.M., S.M., A.Radhakrishnan, A.Rendon, R.R.-S., H.Schepers, J.Sehmi, N.S., H.H.W.S., S.T., T.T., N.V., K.V., P.V., J.Y., W.Z.

Author Information Summary statistics from the genome-wide association study are available from the European Genome-Phenome Archive (EGA, <http://www.ebi.ac.uk/ega/>) under accession number EGAS00000000132. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.C.C. (john.chambers@ic.ac.uk), C.G. (christian.gieger@helmholtz-muenchen.de), P.v.d.H. (p.van.der.harst@umcg.nl), J.S.K. (j.kooner@ic.ac.uk), W.H.O. (who1000@cam.ac.uk) and N.S. (ns6@sanger.ac.uk).

Pim van der Harst^{1,2*}, Weihua Zhang^{3,4*}, Irene Mateo Leach^{1*}, Augusto Rendon^{5,6,7,8*}, Niek Verweij^{1*}, Joban Sehmi^{4,9*}, Dirk S. Paul^{10*}, Ulrich Elling^{11*}, Hooman Allayee¹², Xinzhong Li^{3,14}, Aparna Radhakrishnan^{5,6,8,10}, Sian-Tsung Tan^{4,9}, Katrin Voss^{5,6,8}, Christian X. Weichenberger¹⁵, Cornelis A. Albers^{5,6,10}, Abtehal Al-Hussani³, Folkert W. Asselbergs^{16,17,18}, Marina Ciullo¹⁹, Fabrice Danjou²⁰, Christian Dina^{21,22,23}, Tõnu Esko^{24,25}, David M. Evans²⁶, Lude Franke², Martin Gögele¹⁵, Jaana Hartiala¹², Micha Hersch^{27,28}, Hilma Holm²⁹, Jouke-Jan Hottenga³⁰, Stavroula

Kanoni¹⁰, Marcus E. Kleber^{31,32}, Vasiliiki Lagou^{33,34}, Claudia Langenberg³⁵, Lorna M. Lopez^{26,37}, Leo-Pekka Lyytikäinen^{38,39}, Olle Melander⁴⁰, Federico Murgia⁴¹, Ilya M. Nolte⁴², Paul F. O'Reilly³, Sandosh Padmanabhan⁴³, Afshin Parsa⁴⁴, Nicola Pirastu⁴⁵, Eleonora Porcu⁴⁶, Laura Portas⁴¹, Inga Prokopenko^{33,34}, Janina S. Ried⁴⁷, So-Young Shin¹⁰, Clara S. Tang⁴⁸, Alexander Teumer⁴⁹, Michela Traglia⁵⁰, Sheila Ulivi⁵¹, Harm-Jan Westra², Jian Yang⁵², Jing Hua Zhao³⁵, Franco Anni²⁰, Abdel Abdellaoui³⁰, Antony Attwood^{5,6,8,10}, Beverley Balkau^{53,54}, Stefania Bandinelli⁵⁵, François Bastardot^{56,57}, Beben Benayamin^{48,58}, Bernhard O. Boehm⁵⁹, William O. Cookson⁹, Debashish Das⁶⁰, Paul I. W. de Bakker^{17,18,61,62}, Rudolf A. de Boer¹, Eco J. C. de Geus³⁰, Marleen H. de Moor³⁰, Maria Dimitriou⁶³, Francisco S. Domingues¹⁵, Angela Döring⁶⁴, Gunnar Engström⁴⁰, Gudmundur Ingi Eyjolfsson⁶⁵, Luigi Ferrucci⁶⁶, Krista Fischer²⁴, Renzo Galanello²⁰, Stephen F. Garner^{5,6,8}, Bernd Genser³¹, Quince D. Gibson^{44,67}, Giorgia Grotto⁴⁵, Daniel Fannar Gudbjartsson²⁹, Sarah E. Harris^{37,68}, Anna-Liisa Hartikainen⁶⁹, Claire E. Hastie⁴³, Bo Hedblad⁴⁰, Thomas Illig^{70,71}, Jennifer Jolley^{5,6,8}, Mika Kähönen^{72,73}, Ido P. Kema⁷⁴, John P. Kemp²⁶, Liming Liang⁷⁵, Heather Lloyd-Jones^{5,6,8}, Ruth J. F. Loos³⁵, Stuart Meacham^{5,6,8,10}, Sarah E. Medland⁴⁸, Christa Meisinger⁷⁶, Yasin Memari^{10,77}, Evelin Mihailov¹⁹, Kathy Miller⁷, Miriam F. Moffatt⁹, Matthias Nauck⁷⁸, Maria Novatchkova¹¹, Teresa Nutile¹⁹, Isleifur Olafsson⁷⁹, Pall T. Onundarson^{80,81}, Debora Parracciani⁸², Brenda W. Pennin^{83,84,85}, Lucia Perseu⁴⁶, Antonio Piga⁸⁶, Giorgio Pistis⁵⁰, Anneli Pouta^{87,88}, Ursula Puc¹¹, Olli Raitakari^{89,90}, Susan M. Ring⁹¹, Antonietta Robino⁴⁵, Daniela Ruggiero¹⁹, Aimo Ruukonen⁹², Aude Saint-Pierre¹⁵, Cinzia Sala³⁰, Andres Salumets^{33,94}, Jennifer Sambrook^{5,6,8}, Hein Schepers^{95,96}, Carsten Oliver Schmidt⁹⁷, Herman H. W. Silljé¹, Rob Sladek⁹⁸, Johannes H. Smith⁸³, John M. Starr^{37,99}, Jonathan Stephens^{5,6,8}, Patrick Sulem²⁹, Toshiko Tanaka⁶⁶, Unnur Thorsteinsdottir^{29,100}, Vinicius Tragante¹⁶, Wiek H. van Gilst¹, L. Joost van Pelt⁷⁴, Dirk J. van Veldhuisen¹, Uwe Völker⁴⁹, John B. Whitfield⁴⁸, Gonneke Willemsen³⁰, Bernhard R. Winkelmann¹⁰¹, Gerald Wirnsberger¹¹, Ale Algra^{17,102}, Francesco Cucca^{46,103}, Adamo Pio d'Adamo⁴⁵, John Danesh¹⁰⁴, Ian J. Deary^{36,37}, Anna F. Dominiczak⁴³, Paul Elliott^{37,105}, Paolo Fortina^{106,107}, Philippe Froguel^{108,109}, Paolo Gasparini⁴⁵, Andreas Greinacher¹⁰, Stanley L. Hazen¹¹, Marjo-Riitta Jarvelin^{3,87,105,112,113}, Kay Tee Khaw¹¹⁴, Terho Lehtimäki^{38,39}, Winfried Maerz^{31,115}, Nicholas G. Martin⁴⁸, Andres Metspalu^{24,25}, Braxton D. Mitchell⁴⁴, Grant W. Montgomery⁴⁸, Carmel Moore¹⁰⁴, Gerjan Navis¹¹⁶, Mario Pirastu⁴¹, Peter P. Pramstaller^{15,117,118}, Ramiro Ramirez-Solis¹⁰, Eric Schadt¹¹⁹, James Scott⁹, Alan R. Shuldiner^{44,120}, George Davey Smith²⁶, J. Gustav Smith^{40,121}, Harold Snieder⁴², Rossella Sorice¹⁹, Tim D. Spector¹²², Kari Stefansson^{29,100}, Michael Stummvoll^{123,124}, W. H. Wilson Tang¹¹¹, Daniela Toniolo^{50,125}, Anke Tönjes^{123,124}, Peter M. Vischer^{37,48,52,58}, Peter Vollenweider^{56,57}, Nicholas J. Wareham³⁵, Bruce H. R. Wolfenbutter¹²⁶, Dorret I. Boekmans³⁰, Jacques S. Beckmann^{27,127}, George V. Dedoussis⁶³, Panos Deloukas¹⁰, Manuel A. Ferreira⁴⁸, Serena Sanna⁴⁶, Manuela Uda⁴⁶, Andrew A. Hicks^{15*}, Josef Martin Penninger^{11*}, Christian Gieger^{47*}, Jaspal S. Kooner^{4,9,128*}, Willem H. Ouwehand^{5,6,8,10*}, Nicole Soranzo^{10*} & John C Chambers^{3,4,14,128*}

¹Department of Cardiology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ²Department of Genetics, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ³Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK. ⁴Ealing Hospital NHS Trust, Middlesex UB1 3HW, UK. ⁵Department of Haematology, University of Cambridge, Cambridge CB2 0XY, UK. ⁶NHS Blood and Transplant, Cambridge CB2 0PT, UK. ⁷MRC Biostatistics Unit, Institute of Public Health, Cambridge CB2 2SR, UK. ⁸NIHR Cambridge Biomedical Research Centre, Cambridge CB2 0QQ, UK. ⁹National Heart and Lung Institute, Imperial College London, London W12 0NN, UK. ¹⁰Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. ¹¹Institute of Molecular Biotechnology of the Austrian Academy of Sciences, 1030 Vienna, Austria. ¹²Department of Preventive Medicine, University of Southern California Keck School of Medicine, Los Angeles, California 90033, USA. ¹³Institute of Clinical Sciences, Imperial College London, London W12 0NN, UK. ¹⁴NIHR Cardiovascular Biomedical Research Unit, Royal Brompton and Harefield NHS Foundation Trust and Imperial College London, London SW3 6NP, UK. ¹⁵Center for Biomedicine, European Academy Bozen/Bolzano (EURAC), 39100 Bolzano, Italy. ¹⁶Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, 3508 Utrecht, The Netherlands. ¹⁷Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3508 Utrecht, The Netherlands. ¹⁸Department of Medical Genetics, Biomedical Genetics, University Medical Center Utrecht, 3508 Utrecht, The Netherlands. ¹⁹Institute of Genetics and Biophysics "Adriano Buzzati-Traverso"-CNR, 80131 Naples, Italy. ²⁰Clinica Pediatrica 2a, Dipartimento di Scienze Biomediche e Biotecnologie - Università di Cagliari, Ospedale Regionale Microcitemie ASL8, 09121 Cagliari, Italy. ²¹Institut National de la Santé et de la Recherche Médicale (INSERM) Unité Mixte de Recherche (UMR) 1087, BP 70721 44007 Nantes cedex, 1 Nantes, France. ²²Centre National de la Recherche Scientifique (CNRS) UMR 6291, BP 70721 44007 Nantes cedex 1, Nantes, France. ²³School of Medicine, Nantes University, 44000 Nantes, France. ²⁴Estonian Genome Center of University of Tartu, 51010 Tartu, Estonia. ²⁵Institute of Molecular and Cell Biology, University of Tartu, 51010 Tartu, Estonia. ²⁶MRC Centre for Causal Analyses in Translational Epidemiology, School of Social and Community Medicine, University of Bristol, Bristol BS8 2BN, UK. ²⁷Department of Medical Genetics, University of Lausanne, CH-1005 Lausanne, Switzerland. ²⁸Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland. ²⁹deCODE genetics, 101 Reykjavik, Iceland. ³⁰Department of Biological Psychology, VU University, 1081 BT Amsterdam, The Netherlands. ³¹Mannheim Institute of Public Health, Social and Preventive Medicine, Medical Faculty of Mannheim, University of Heidelberg, D-68167 Mannheim, Germany. ³²LURIC Study nonprofit LLC, D-79098 Freiburg, Germany. ³³Oxford Centre for Diabetes, Endocrinology and Metabolism, Radcliffe Department of Medicine, University of Oxford OX3 7JL, UK. ³⁴Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. ³⁵MRC Epidemiology Unit, Institute of Metabolic Science, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK. ³⁶Department of Psychology, The University of Edinburgh, Edinburgh EH8

- 9JZ, UK. ³⁷Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh EH8 9JZ, UK. ³⁸Department of Clinical Chemistry, Fimlab Laboratories, Tampere University Hospital, FIN-33521 Tampere, Finland. ³⁹Department of Clinical Chemistry, University of Tampere School of Medicine, FIN-33521 Tampere, Finland. ⁴⁰Department of Clinical Sciences, Lund University, SE-205 02 Malmö, Sweden. ⁴¹Institute of Population Genetics, National Research Council of Italy, 07100 Sassari, Italy. ⁴²Department of Epidemiology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ⁴³Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow G12 8QQ, United Kingdom. ⁴⁴University of Maryland School of Medicine, Baltimore, Maryland 21201, USA. ⁴⁵Institute for Maternal and Child Health-IRCCS "Burlo Garofolo"-Trieste, University of Trieste, 34137 Trieste, Italy. ⁴⁶Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche, c/o Cittadella Universitaria di Monserrato, Monserrato, Cagliari 09042, Italy. ⁴⁷Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany. ⁴⁸Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia. ⁴⁹Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, D-17487 Greifswald, Germany. ⁵⁰Division of Genetics and Cell Biology, San Raffaele Scientific Institute, 20132 Milano, Italy. ⁵¹Institute for Maternal and Child Health - IRCCS "Burlo Garofolo"-Trieste, 34137 Trieste, Italy. ⁵²University of Queensland Diamantina Institute, The University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4102, Australia. ⁵³Inserm, CESP Centre for research in Epidemiology and Population Health, U1018, Villejuif F-94807, France. ⁵⁴University Paris Sud 11, UMRs 1018, Villejuif F-94807, France. ⁵⁵Geriatric Unit, Azienda Sanitaria Firenze, 50125 Florence, Italy. ⁵⁶Centre Hospitalier Universitaire Vaudois, CH-1011 Lausanne, Switzerland. ⁵⁷Department of Internal Medicine, University of Lausanne, CH-1011 Lausanne, Switzerland. ⁵⁸Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia. ⁵⁹Division of Endocrinology and Diabetes, Department of Medicine, University Hospital, Ulm D-89075, Germany. ⁶⁰The Hatter Cardiovascular Institute, University College London, London WC1E 6HX, UK. ⁶¹Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁶²Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02142, USA. ⁶³Nutrition and Dietetics, Harokopio University, Kallithea 17671, Athens, Greece. ⁶⁴Institute of Epidemiology I and Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany. ⁶⁵The Laboratory in Mjodd, 108 Reykjavik, Iceland. ⁶⁶Clinical Research Branch, National Institute on Aging, Baltimore, Maryland 21250, USA. ⁶⁷Instituto de Saúde Coletiva, Federal University of Bahia, Salvador, Bahia 40110-040, Brazil. ⁶⁸Medical Genetics Section, The University of Edinburgh, Edinburgh EH4 2XU, UK. ⁶⁹Institute of Clinical Sciences, Obstetrics and Gynecology, University of Oulu FIN-90220 Oulu, Finland. ⁷⁰Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany. ⁷¹Hannover Unified Biobank, Hannover Medical School, D-30625 Hannover, Germany. ⁷²Department of Clinical Physiology, Tampere University Hospital, FIN-33521 Tampere, Finland. ⁷³Department of Clinical Physiology, University of Tampere School of Medicine, FIN-33521 Tampere, Finland. ⁷⁴Department of Laboratory Medicine, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ⁷⁵Department of Epidemiology, Department of Biostatistics, Harvard School of Public Health, Cambridge, Massachusetts 02115, USA. ⁷⁶Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, D-85764 Neuherberg, Germany. ⁷⁷Department of Twin Research and Genetic Epidemiology, Kings College London, London SE1 7EH, UK. ⁷⁸Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, D-17475 Greifswald, Germany. ⁷⁹Department of Clinical Biochemistry, Landspítali University Hospital, 101 Reykjavik, Iceland. ⁸⁰Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland. ⁸¹Laboratory of Hematology and Coagulation Disorder Center, Landspítali University Hospital, 101 Reykjavik, Iceland. ⁸²Genetic Park of Ogliastra, Perdasdefogu, Sardinia, Italy. ⁸³Department of Psychiatry/EMGO Institute/Neuroscience Campus, VU University Medical Centre, 1081 BT Amsterdam, The Netherlands. ⁸⁴Department of Psychiatry, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ⁸⁵Department of Psychiatry, Leiden University Medical Centre, 2333 Leiden, The Netherlands. ⁸⁶Division of Pediatrics and Thalassemia Centre, Department of Clinical and Biological Sciences, University of Torino, 10043 Orbassano, Turin, Italy. ⁸⁷Institute of Health Sciences, University of Oulu, FIN-90220 Oulu, Finland. ⁸⁸National Institute of Health and Welfare, Aapistie 1, P.O. Box 310, FIN-90101 Oulu, Finland. ⁸⁹Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, FIN-20521 Turku, Finland. ⁹⁰Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, FIN-20521 Turku, Finland. ⁹¹The School of Social and Community Medicine, University of Bristol, Bristol BS8 2PS, UK. ⁹²Institute of Diagnostics, University of Oulu, FIN-90014 Oulu, Finland. ⁹³Competence Centre on Reproductive Medicine and Biology, 50410 Tartu, Estonia. ⁹⁴Institute of General and Molecular Pathology, University of Tartu, 51014 Tartu, Estonia. ⁹⁵Department of Experimental Hematology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ⁹⁶Department of Stem Cell Biology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ⁹⁷Institute for Community Medicine, University Medicine Greifswald, D-17475 Greifswald, Germany. ⁹⁸Departments of Human Genetics and Medicine, Faculty of Medicine, McGill University, Montreal, Quebec H3A 1B1, Canada. ⁹⁹Geriatric Medicine Unit, The University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK. ¹⁰⁰Faculty of Medicine, University of Iceland, 101 Reykjavik, Iceland. ¹⁰¹ClinPhenomics Study Center, D-60594 Frankfurt, Germany. ¹⁰²Utrecht Stroke Center, Department of Neurology and Neurosurgery, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, 3584 CG Utrecht, The Netherlands. ¹⁰³Dipartimento di Scienze Biomediche, Università di Sassari, 07100 Sassari, Italy. ¹⁰⁴Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. ¹⁰⁵MRC-HPA Centre for Environment and Health, Imperial College London, London W2 1PG, UK. ¹⁰⁶Department of Cancer Biology, Thomas Jefferson University Jefferson Medical College, Philadelphia, Pennsylvania 19107, USA. ¹⁰⁷Dipartimento di Medicina Molecolare, Università La Sapienza, 00161 Roma, Italy. ¹⁰⁸Centre National de la Recherche Scientifique (CNRS)-Unité mixte de recherche (UMR) 8199, Lille Pasteur Institute, Lille 59100, France. ¹⁰⁹Department of Genomics of Common Disease, School of Public Health, Imperial College London, London W2 1PG, UK. ¹¹⁰Institute for Immunology and Transfusion Medicine, University Medicine Greifswald, D-17487 Greifswald, Germany. ¹¹¹Center for Cardiovascular Diagnostics and Prevention, Department of Cell Biology, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio 44195, USA. ¹¹²Department of Lifecourse and Service, National Institute for Health and Welfare, FIN-90101 Oulu, Finland. ¹¹³Biocenter Oulu, University of Oulu, FIN-90220 Oulu, Finland. ¹¹⁴Clinical Gerontology Unit, Box 251, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK. ¹¹⁵Synlab Academy, D-68165 Mannheim, Germany. ¹¹⁶Department of Internal Medicine, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ¹¹⁷Department of Neurology, General Central Hospital, 39100 Bolzano, Italy. ¹¹⁸Department of Neurology, University of Lübeck, D-23538 Lübeck, Germany. ¹¹⁹Institute for Genomics and Multiscale Biology, Mount Sinai School of Medicine, New York 10029-6574, USA. ¹²⁰Geriatric Research and Education Clinical Center, Veterans Administration Medical Center, Baltimore, Maryland 21201, USA. ¹²¹Department of Cardiology, Lund University, 22185 Lund, Sweden. ¹²²Department of Twin Research and Genetic Epidemiology, Kings College London, London SE1 7EH, UK. ¹²³Department of Medicine, University of Leipzig, Liebigstr. 18, D-04103 Leipzig, Germany. ¹²⁴University of Leipzig, IFB Adiposity Diseases, D-04103 Leipzig, Germany. ¹²⁵Institute of Molecular Genetics, CNR, 27100 Pavia, Italy. ¹²⁶Department of Endocrinology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. ¹²⁷Service of Medical Genetics, Centre Hospitalier Universitaire Vaudois, 1011 Lausanne, Switzerland. ¹²⁸Imperial College Healthcare NHS Trust, London W12 0HS, UK.

*These authors contributed equally to this work.

METHODS

Genome-wide association. Genome-wide association was carried out in 62,553 people of European ancestry and 9,308 people of South Asian ancestry, using up to 2,644,161 autosomal and 67,645 X-chromosome SNPs. Imputation was done using haplotypes from HapMap Phase 2. Characteristics of participants, genotyping arrays and imputation are summarized in Supplementary Tables 1 and 2. Participants with extreme measurements ($> \pm 3$ s.d. from mean) were excluded on a per-phenotype basis. Each population cohort was approved by a research ethics committee, and all participants gave informed consent.

SNP associations with each phenotype were tested by linear regression using an additive genetic model. Associations were tested separately in men and women in each cohort, with principal components and other study-specific factors as covariates to account of population substructure as described in Supplementary Table 2. Test statistics from each cohort were then corrected for their respective genomic-control inflation factor to adjust for residual population sub-structure; genomic-control inflation factors are summarized in Supplementary Table 3. We then carried out a meta-analysis of results from the individual cohorts using Z-scores weighted by the square root of sample size. The meta-analysis was varied out among Europeans and South Asians separately. There were no South-Asian-specific discoveries, but also little evidence for heterogeneity of effect at known or new genetic loci (Supplementary Table 23); we therefore carried out a final combined analysis of results for the two populations. SNPs with minor allele frequency $<1\%$ (weighted average across cohorts) were removed, as were SNPs with weight $<50\%$ of phenotype sample size. There was no evidence for inflation of test statistics at SNPs not known to be associated with red blood cell phenotypes (Supplementary Table 3), and genomic control was not applied to the final meta-analysis results. We used the function 'clump' implemented in PLINK to cluster the SNPs into genomic loci using a 2-Mb window; clustering was done separately for each phenotype. Inverse variance meta-analysis was used to quantify effect sizes for SNPs of interest.

Genome-wide significance was inferred at $P < 1 \times 10^{-8}$. This choice of statistical threshold was grounded on the guidelines derived from studies of the ENCODE (encyclopedia of DNA elements) regions⁶, combined with results of permutation testing to determine the additional adjustment needed for the six red blood cell phenotypes studied (Supplementary Tables 24, 25 and Supplementary Note). As an alternative strategy, a P -value threshold of $P < 3.2 \times 10^{-9}$ would provide correction for the number of SNP-phenotype combinations tested without any adjustment for the correlations between the SNPs or phenotypes tested. We note that 70 of the 75 loci identified would exceed such a highly stringent threshold, including all four of the loci identified through the joint analysis of European and South Asian data.

Replication testing. We carried out replication testing of 22 SNPs selected on the basis of the following criteria: (1) the lead SNP from each of 17 loci showing suggestive evidence for association with one or more red blood cell phenotypes in Europeans ($P > 10^{-8}$ and $P < 10^{-7}$), and (2) the lead SNP from each of the loci identified through combined analysis of genome-wide association data for Europeans and South Asians. Replication testing was done using a combination of *in silico* results and direct genotyping among 63,506 people from four population cohorts.

In silico data were available for 34,843 people from Iceland participating in the deCODE (diabetes epidemiology: collaborative analysis of diagnostic criteria in Europe) study³⁷ (Supplementary Table 1). SNPs were directly genotyped with the Illumina HumanHap300 or CNV370 chips or imputed from one or more of four sources: the HapMap2 CEU sample (60 triads), the 1000 Genomes Project data (179 individuals) and Icelandic samples genotyped with the Illumina Human1 M-Duo (123 triads) or the HumanOmni1-Quad chips (505 individuals), as previously described in ref. 37. The 22 SNPs were tested for association against their respective discovery phenotypes, under an additive genetic model; results were combined with the genome-wide association data by weighted-Z-score meta-analysis.

We found that for 7 of the 22 SNPs carried forward for replication, their associations with phenotype remained inconclusive after *in silico* testing ($P > 10^{-8}$ but $P < 10^{-7}$). For these SNPs we carried out additional direct genotyping using Sequenom assays, among up to 20,066 people from three population cohorts (Supplementary Table 1). Associations were tested in each cohort separately, and results combined across the replication cohorts, and then with the genome-wide association data, by weighted-Z-score meta-analysis (Supplementary Table 26).

Conditional analysis. We performed conditional-association analysis using the summary statistics from the meta-analysis to test for the association of each SNP while conditioning on the top SNPs, with correlations between SNPs due to linkage disequilibrium estimated from the imputed genotype data from the atherosclerosis risk in communities (ARIC) cohort^{8,38}. Secondary-association signals were selected with conditional-association $P < 1 \times 10^{-8}$.

Identification of candidate genes. We considered the nearest gene, and any other gene located within 10 kb of the sentinel SNP, to be a candidate for mediating the association with red blood cell phenotype. We also used coding variant, eQTL and literature analyses to identify candidate genes. On the basis of analysis of linkage-disequilibrium relations at the 75 genetic loci, we defined genomic region as the 1-Mb interval either side of the sentinel SNP for our functional genomic studies (Supplementary Fig. 7).

Coding variation. We identified all non-synonymous SNPs that were in linkage disequilibrium with one or more of the sentinel SNPs at $r^2 > 0.8$ in 1000 Genomes Project data set (released in March 2012). We considered the gene to be a candidate when the non-synonymous and sentinel SNPs were in linkage disequilibrium at $r^2 > 0.8$ and with no evidence for heterogeneity of effect on phenotype. This strategy identified 39 non-synonymous SNPs distributed between 24 genes (Supplementary Table 9), representing a \sim sixfold enrichment compared to the mean number expected under the null hypothesis generated by permutation testing of SNP sets matched for allele frequency (± 0.05) and number of genes in proximity (± 10 kb), but selected otherwise at random ($P = 0.01$; Supplementary Note).

Expression analyses. To identify the possible genes influencing red blood cell phenotypes at the 75 loci, we examined the association of the sentinel SNPs with eQTL data from two data sets: (1) peripheral blood lymphocytes from 206 families of European descent (830 parents and offspring)³⁹ and (2) peripheral blood lymphocytes from 1,469 unrelated individuals⁴⁰.

SNPs were tested for association with expression of nearby (1 Mb) genes ($P < 0.05$ after Bonferroni correction for number of SNP-transcript associations tested). Where eQTLs were identified, we used the whole-genome SNP data available in these data sets (imputed with HapMap Phase 2 genotypes), to identify the SNP at the locus most closely associated with transcript level (the transcript SNP). We then tested whether the sentinel SNP and the transcript SNP were coincident, defined as $r^2 > 0.8$ with no evidence for heterogeneity of effect on phenotype or transcript level ($P > 0.05$). This strategy identified eQTLs involving 28 genes from 18 loci (Supplementary Table 10).

GRAIL analyses. We carried out a literature analysis using the GRAIL algorithm⁹, a statistical tool that uses text mining of PubMed abstracts to annotate candidate genes from loci associated with phenotypic traits. We carried out the analysis using the 2006 data set to avoid confounding by subsequent GWAS discoveries; results identified candidate genes at nine loci ($P < 0.05$; Supplementary Table 11). Results are also shown for a GRAIL analysis using the 2011 PubMed data set, although these were not used for the final analysis.

Gene expression in haematopoietic precursors. Cord-blood-derived CD34⁺ haematopoietic stem cells were differentiated *in vitro* along the erythroid lineage in the presence of 6 U ml⁻¹ erythropoietin (R&D Systems), 10 ng ml⁻¹ interleukin (IL)-3 (Miltenyi Biotec) and 100 ng ml⁻¹ stem cell factor (R&D Systems). Cells were collected at days 3, 5, 7, 9 and 10 in three biological replicates and gene expression was assayed using Illumina human WGv3.0 microarrays⁴¹. For each gene, we determined the relationship of gene expression with time using linear regression, and calculated the t -statistic for the difference in β from zero. We then classified gene-expression patterns as increasing, decreasing or unchanged on the basis of the 2.5% and 97.5% quartiles of the t distribution with 4 degrees of freedom. To test whether a gene set was enriched for differentially regulated genes, a Wilcoxon signed-rank test of the t scores in the gene set relative to all others genes that were expressed in at least one time point was calculated.

FAIRE-seq. We generated maps of chromatin accessibility ('open chromatin') in primary human erythroblasts and megakaryocytes, and in peripheral blood monocytes using FAIRE-seq. Cord-blood-derived CD34⁺ haematopoietic progenitor cells from two unrelated individuals were differentiated *in vitro* into either erythroblasts (in the presence of erythropoietin, IL-3 and stem cell factor) or megakaryocytes (in the presence of thrombopoietin and IL-1 β). Monocytes were purified from leukocyte cones of apheresis collections from another two individuals.

FAIRE experiments were performed as previously described in ref. 42. FAIRE DNA was processed following the Illumina paired-end library-generation protocol. Genomic libraries derived from erythroblast and megakaryocyte cultures were sequenced with 54-bp paired-end reads on Illumina Genome Analyzer II. Libraries derived from monocyte extractions were sequenced with 50-bp paired-end reads on Illumina HiSeq. Raw sequence reads were aligned to the human reference sequence (NCBI build 37) using the read mapper Stampy⁴³. Reads were realigned around known insertions and deletions, followed by base-quality recalibration using the Genome Analysis Toolkit (GATK)⁴⁴. Duplicates were flagged using the software Picard (<http://picard.sourceforge.net/>) and excluded from subsequent analyses. For each cell type, we merged all read fragments into one data set. NDRs were identified as regions of sequencing enrichment (peaks) using the software F-Seq³⁶. We applied a feature length of $L = 600$ bp and a s.d. threshold of $T = 8.0$ over the mean across a local background. In order to reduce

false-positive peak calls, we removed regions of collapsed repeats as recently described, applying a threshold of 0.1%⁴⁵. For each associated locus, candidate functional SNPs were selected by identifying all biallelic SNPs with an $r^2 > 0.8$ and within 1 Mb of the sentinel SNP in the European samples of the 1000 Genomes Project (data released June 2011).

***D. melanogaster* gene-silencing models.** We used haemocyte-specific RNAi silencing to investigate whether the 121 candidate genes identified in the red blood cell GWAS influenced blood cell formation in *D. melanogaster*. We identified *D. melanogaster* genes predicted to be orthologues of human genes using the Ensembl v65 Compara pipeline, an established phylogenetic-tree-based approach for orthology prediction⁴⁶; this revealed 96 *D. melanogaster* orthologues for 74 of the 121 human candidate genes (Supplementary Table 27). We evaluated each of the 96 orthologues for a blood cell phenotype in *D. melanogaster*. We obtained all 225 available *D. melanogaster* lines carrying inducible siRNA constructs from the Vienna *Drosophila* RNAi Center (VDRC)²³. To achieve haemocyte-specific knockdowns, flies were crossed to the blood-specific *Hml-Gal4* line driving Gal4 expression under the control of a hemolectin promoter⁴⁷. Flies were crossed at 29 °C, and early and late L3 larvae analysed 7 days after mating. Upstream activating sequence–green fluorescent protein enabled microscopic visualization of plasmatocytes and evaluation of cell size and cell number (L3 larvae only). Early- and late-stage larvae were incubated at 60 °C for 15 min, a process that turns the crystal cells black and allows quantification of crystal cells microscopically. For each orthologue, all available RNAi silencer constructs were investigated, and in addition, each construct was assayed in duplicate, blind to initial result. Cell counts were quantified visually (0–3, decreased or increased) and the mean of the duplicate measurements calculated.

We separately carried out permutation testing in a genome-wide screen of 5,658 *D. melanogaster* genes to simulate expectations under the null hypothesis (Supplementary Fig. 8 and Supplementary Note); results confirmed that the 121 candidate genes were enriched for blood cell phenotype in *D. melanogaster* orthologues ($P < 0.05$), and showed that this was robust to threshold for calling. **Contribution of the genetic loci identified to population variation in red blood cell phenotypes.** This was investigated in participants from the Estonian Genome Center of University of Tartu (EGCUT), LIFELINES, Ludwigshafen Risk and Cardiovascular Health Study (LURIC) and Young Finns cohorts using samples that were not included in the discovery experiment (Supplementary Table 1). The contribution of the SNPs to population variation in red blood cell phenotypes was quantified using two models: model 1, limited to SNPs associated

with respective phenotype at $P < 1 \times 10^{-8}$; and model 2, comprising all of the 75 sentinel SNPs identified. Estimates of population variance explained were made in each study separately, and average values calculated weighted by sample size (Supplementary Table 21).

We then investigated whether the 75 sentinel SNPs influenced the probability of being in the highest versus the lowest quartile for population distribution of phenotype. Two SNP scores were calculated for each phenotype: score 1, limited to SNPs associated with respective phenotype at $P < 1 \times 10^{-8}$, and score 2, containing all 75 sentinel SNPs identified. For both, SNP score was calculated as the sum of number of effect (trait raising) alleles present, weighted according to effect size. We then calculated the odds ratio for being in the highest versus the lowest quartile of phenotype, associated with SNP scores in the second, third and fourth quartiles, compared to first quartile of SNP score. Odds ratios were calculated in each study separately, and then combined by inverse variance meta-analysis (Fig. 3).

37. Holm, H. *et al.* A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nature Genet.* **43**, 316–320 (2011).
38. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
39. Dixon, A. L. *et al.* A genome-wide association study of global gene expression. *Nature Genet.* **39**, 1202–1207 (2007).
40. Dubois, P. C. *et al.* Multiple common variants for celiac disease influencing immune gene expression. *Nature Genet.* **42**, 295–302 (2010).
41. Anderson, R. J. *et al.* Reduced dependency on arteriography for penetrating extremity trauma: influence of wound location and noninvasive vascular studies. *J. Trauma* **30**, 1059–1063 (1990).
42. Giresi, P. G. & Lieb, J. D. Isolation of active regulatory elements from eukaryotic chromatin using FAIRE (Formaldehyde Assisted Isolation of Regulatory Elements). *Methods* **48**, 233–239 (2009).
43. Lunter, G. & Goodson, M. Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome Res.* **21**, 936–939 (2011).
44. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
45. Pickrell, J. K., Gaffney, D. J., Gilad, Y. & Pritchard, J. K. False positive peaks in ChIP-seq and other sequencing-based functional assays caused by unannotated high copy number regions. *Bioinformatics* **27**, 2144–2146 (2011).
46. Vilella, A. J. *et al.* EnsemblCompara GeneTrees: complete, duplication-aware phylogenetic trees in vertebrates. *Genome Res.* **19**, 327–335 (2009).
47. Goto, A. *et al.* A *Drosophila* haemocyte-specific protein, hemolectin, similar to human von Willebrand factor. *Biochem. J.* **359**, 99–108 (2001).

Functional screening identifies miRNAs inducing cardiac regeneration

Ana Eulalio^{1†}, Miguel Mano¹, Matteo Dal Ferro^{1,2}, Lorena Zentilin¹, Gianfranco Sinagra², Serena Zacchigna¹ & Mauro Giacca¹

In mammals, enlargement of the heart during embryonic development is primarily dependent on the increase in cardiomyocyte numbers. Shortly after birth, however, cardiomyocytes stop proliferating and further growth of the myocardium occurs through hypertrophic enlargement of the existing myocytes. As a consequence of the minimal renewal of cardiomyocytes during adult life, repair of cardiac damage through myocardial regeneration is very limited. Here we show that the exogenous administration of selected microRNAs (miRNAs) markedly stimulates cardiomyocyte proliferation and promotes cardiac repair. We performed a high-content microscopy, high-throughput functional screening for human miRNAs that promoted neonatal cardiomyocyte proliferation using a whole-genome miRNA library. Forty miRNAs strongly increased both DNA synthesis and cytokinesis in neonatal mouse and rat cardiomyocytes. Two of these miRNAs (hsa-miR-590 and hsa-miR-199a) were further selected for testing and were shown to promote cell cycle re-entry of adult cardiomyocytes *ex vivo* and to promote cardiomyocyte proliferation in both neonatal and adult animals. After myocardial infarction in mice, these miRNAs stimulated marked cardiac regeneration and almost complete recovery of cardiac functional parameters. The miRNAs identified hold great promise for the treatment of cardiac pathologies consequent to cardiomyocyte loss.

In mammals, cardiomyocyte (CM) proliferation rapidly ceases after birth^{1–3}. The transition of CMs from a proliferative state, characteristic of embryonic stages, to the differentiated, hypertrophic phenotype typical of adult cells is a highly regulated process^{1–3}. Although there is also evidence of cardiac cell renewal in humans^{4,5}, the proliferative capacity of adult CMs remains limited. Consequently, the ability of the adult heart to repair itself following injury, such as myocardial infarction or heart failure, is very restricted. A number of reports have indicated that, after cardiac injury, a few replicating CMs can be detected in the damaged area, suggestive of a partial attempt at myocardial regeneration^{6–10}; however, repair typically occurs through a scarring mechanism.

Notably, CMs isolated from neonatal hearts retain some proliferative capacity in culture, which becomes extinguished after a few days³. This offers a valuable opportunity to identify the mechanisms that maintain this process, with the goal to foster CM proliferation and induce cardiac regeneration after injury.

MicroRNAs (miRNAs) regulate gene expression post-transcriptionally, by base-pairing to partially complementary sequences in target messenger RNAs^{11,12}. Impairment of the miRNA pathway in cardiac muscle leads to heart failure and cardiomyopathy^{13,14}. Furthermore, altered miRNA expression patterns have been associated with various cardiac pathologies^{15–18}. Although some miRNAs have been shown to inhibit CM proliferation, including miR-1^{19,20}, miR-133²¹ and members of the miR-15 family²², to date and to our knowledge no miRNA has been reported to increase CM proliferation.

Here, we set out to systematically identify miRNAs triggering CM proliferation using a synthetic miRNA library of human origin, with the ultimate purpose to identify potential human therapeutics.

Screening for miRNAs regulating CM proliferation

We performed a high-content, fluorescence-microscopy-based, high-throughput screening in neonatal rat CMs using a library of 875

miRNA mimics (988 mature miRNAs, 875 unique sequences, miRBase release 13.0 (2009), <http://mirbase.org>; Fig. 1a). Cultures of neonatal rat ventricular CMs, containing >90% CMs, were transfected with the library of miRNA mimics (>95% transfection efficiency; Supplementary Fig. 1a, b). After 72 h, the cells were stained for sarcomeric α -actinin to distinguish CMs, for the proliferation antigen Ki-67 and for 5-ethynyl-2'-deoxyuridine (EdU), a uridine analogue that is incorporated into newly synthesized DNA²³ (Fig. 1b). There was a high correlation between the two proliferation markers (Spearman $r > 0.80$; Fig. 1b and Supplementary Fig. 1c). Image segmentation and analysis was performed to selectively quantify the number of proliferating CMs (α -actinin⁺, Ki-67⁺, EdU⁺; Fig. 1b). The screening was performed in duplicate; the replicates showed very good reproducibility (Spearman $r = 0.94$; Fig. 1c). On average, approximately 2,500 cells were analysed per experimental condition and replicate.

The screening identified 204 miRNAs that significantly increased neonatal CM proliferation more than twofold when compared to CMs in basal conditions or treated with the control miRNA cel-miR-67 (from 12.5% up to more than 45% proliferating cells; red in Fig. 1c and Supplementary Table 1); 331 miRNAs decreased EdU incorporation and Ki-67 positivity (from 12.5% to 0%; blue in Fig. 1c, Supplementary Fig. 2 and Supplementary Table 1), the majority without significantly affecting cell viability. Examples of miRNAs strongly increasing or decreasing CM proliferation are shown in Fig. 1d. The majority of the miRNAs that enhanced EdU and Ki-67 positivity also increased the number of CMs (Spearman $r = 0.63$; Fig. 1e). Of note, miRNAs belonging to the same family induced a matched proliferation phenotype (Fig. 1f). No significant synergistic effect was observed upon pairwise administration of the top four miRNAs (Supplementary Fig. 3). Comparison of miRNA expression profiles in neonatal and adult CMs showed that, out of the

¹Molecular Medicine Laboratory, International Centre for Genetic Engineering and Biotechnology (ICGEB), 34149 Trieste, Italy. ²Department of Medical, Surgical and Health Sciences, University of Trieste, Trieste, Italy and Center for Translational Cardiology, Azienda Ospedaliero-Universitaria "Ospedali Riuniti di Trieste", 34129 Trieste, Italy. [†]Present address: Institute for Molecular Infection Biology (IMIB), University of Würzburg, D-97080 Würzburg, Germany.

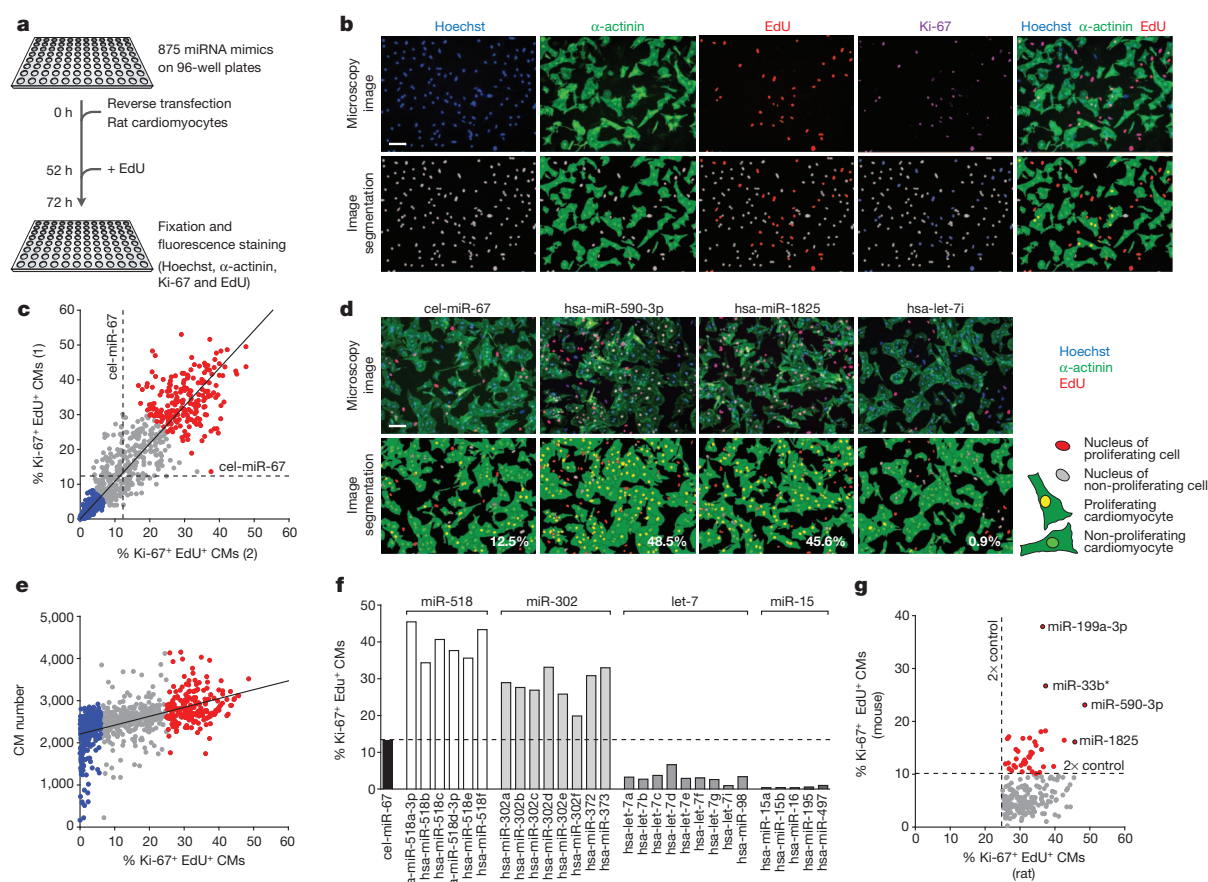


Figure 1 | High-content screening identifies miRNAs regulating CM proliferation. **a**, Screening workflow. **b**, Microscopy images and image reconstruction of rat CMs treated with cel-miR-67 (control). **c**, Percentage of proliferating rat CMs following treatment with the 875 miRNAs. Red, >twofold proliferation; blue, <twofold proliferation. Approximately 2,500 cells were analysed per miRNA/replicate. **d**, Rat CMs treated with selected miRNAs.

269 miRNAs expressed by rat neonatal CMs, the majority of the pro-proliferative miRNAs were less expressed in adult cells (Supplementary Fig. 4).

Because the screening was based on high-content imaging, we also determined the effect of the 875 miRNAs on CM size. The top miRNAs promoting CM proliferation were not among those inducing hypertrophy (for example, hsa-miR-29b; Supplementary Fig. 5a, b), as also confirmed by measuring the atrial natriuretic peptide levels (Supplementary Fig. 5c–f).

To understand whether the identified miRNAs also increased proliferation of CMs from another species, thus increasing the likelihood of the conservation of their functional effect in human cells, we re-tested the selected 204 miRNAs in neonatal mouse CMs. In the mouse heart, CMs stop dividing sooner than those in the rat heart (shortly after birth in the mouse and 3–4 days after birth in the rat^{24,25}). Consequently, CMs isolated from newborn mice have a lower proliferative capacity (approximately 5%). From the 204 miRNAs tested, 40 miRNAs also enhanced EdU incorporation and Ki-67 positivity in mouse CMs at least twofold (Fig. 1g, Supplementary Table 1 and Supplementary Fig. 6).

To demonstrate that the increase in CM DNA replication eventually resulted in an increase in karyokinesis and cytokinesis, we treated rat CMs with the top 10 miRNAs from the rat screening and mouse CMs with the top 10 from the mouse re-screening. After 72 h, the cells were stained for histone H3 phosphorylated on serine 10 (H3S10ph), a marker of late G2/mitosis, and for Aurora B kinase localization in midbodies, transient structures formed during

cytokinesis (Fig. 2a, b). Treatment with the selected miRNAs significantly increased both the percentage of cells positive for H3S10ph, and the number of cells presenting midbodies. Table 1 reports the results of these quantifications. Notably, all the 10 plus 10 selected miRNAs were effective in promoting karyokinesis and cell division in both rat and mouse cells. Figure 2c, d shows the results for the best two performers of each list. Consistent with an effective cell division, at 6 days after transfection we observed a significant increase in CM number (Fig. 2e and Supplementary Fig. 7). None of the top performing miRNAs increased rat cardiac fibroblast proliferation (Supplementary Fig. 8a, b), indicating that their effect for CMs was rather selective.

miRNAs increase post-natal CM proliferation

Next we tested whether the identified miRNAs stimulated post-natal CM proliferation. As a first approach, we used CMs isolated from 7-days-old rats. At this age, CM proliferation was almost undetectable; however, the percentage of EdU⁺ CMs significantly increased up to 13–20% upon treatment with hsa-miR-590-3p, hsa-miR-199a-3p, hsa-miR-1825 and hsa-miR-33b* (Supplementary Fig. 9).

To test whether these miRNAs also induce re-entry into the cell cycle of fully differentiated cells, CMs were isolated from adult (2-months old) rats and transfected with cel-miR-67 or with hsa-miR-590-3p and hsa-miR-199a-3p. Remarkably, treatment with the two miRNAs determined a time-dependent re-entry of the cells into the cell cycle (Fig. 2f–h), eventually leading to a significant increase of the number of CMs (Fig. 2h).

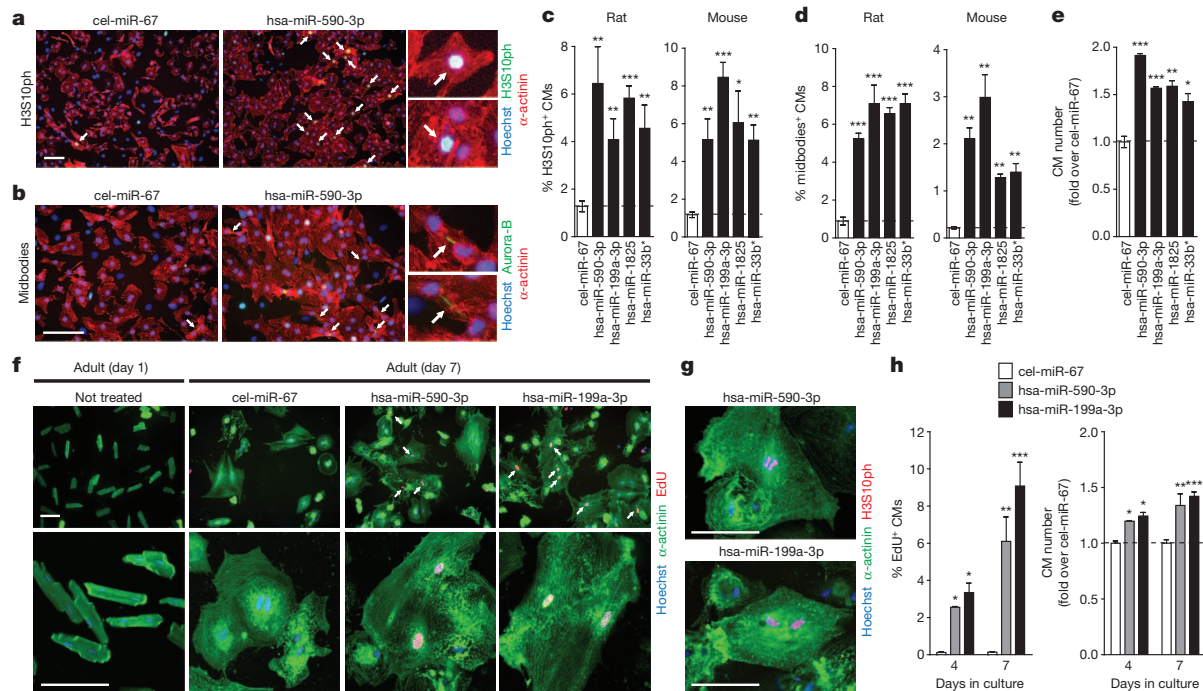


Figure 2 | miRNAs increase cytokinesis and proliferation of post-natal CMs. **a, b**, H3S10ph (**a**) and Aurora B kinase (**b**) staining of miRNA-treated rat CMs. Arrows, positive CMs. **c, d**, Percentage of CMs positive for H3S10ph (**c**) and midbodies (**d**). $n = 3-5$ per group. **e**, CM number, 6 days after miRNA treatment. $n = 3$ per group. **f**, EdU staining of adult rat CMs. Top,

epifluorescence; bottom, confocal. Arrows, proliferating CMs. **g**, H3S10ph staining of miRNA-treated adult rat CMs, at day 7. **h**, Percentage of proliferating adult rat CMs and CM number. $n = 3-6$ per group. All panels, mean \pm s.e.m.; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ relative to cel-miR-67; scale bars, 100 μ m.

To obtain further molecular insights into the capacity of the miRNAs to induce cell cycle re-entry of adult CMs, we compared neonatal, adult CMs at the time of isolation and adult CMs transfected with cel-miR-67, hsa-miR-590-3p and hsa-miR-199a-3p at day 7 for the expression of genes related with cardiac differentiation²⁶ and proliferation. CM culture itself, as described previously²⁷⁻²⁹, determined a partial de-differentiation of adult CMs, as concluded from increased expression of β -myosin heavy chain (MYH7), decrease in α -myosin heavy chain (MYH6), and increase in DAB2, destrin and RUNX1 (Supplementary Fig. 10 and Fig. 2f for microscopy images showing

more rounded cells, with less organized contractile apparatus); these changes in gene expression never reached those found in of neonatal CMs, except for NKX2-5. This partial de-differentiation, however, did not result in cell cycle re-entry per se, unless the cells were stimulated with hsa-miR-590-3p and hsa-miR-199a-3p. Upon miRNA treatment, the marked increase in the levels of cyclin genes correlated with their increased proliferation (Supplementary Fig. 10 and Fig. 2f-h).

Collectively, these experiments demonstrate that hsa-miR-590-3p and hsa-miR-199a-3p are able to induce proliferation of post-natal CMs.

Table 1 | Effect of selected miRNAs on rat and mouse CMs.

	miRNA	Proliferating CMs (%)†	H3S10ph ⁺ CMs (%)†	Midbodies ⁺ CMs (%)†
Rat CMs	cel-miR-67	11.1 \pm 1.2	1.28 \pm 0.23	0.89 \pm 0.20
	hsa-miR-590-3p	49.1 \pm 3.5	6.44 \pm 1.55	5.23 \pm 0.30
	hsa-miR-1825	48.2 \pm 4.2	5.83 \pm 0.51	6.56 \pm 0.32
	hsa-miR-518a-3p	46.9 \pm 1.0	6.48 \pm 0.57	4.57 \pm 0.04
	hsa-miR-1273	41.8 \pm 2.2	5.47 \pm 0.20	3.41 \pm 0.15
	hsa-miR-302b*	43.0 \pm 1.3	6.53 \pm 0.70	3.02 \pm 0.14
	hsa-miR-495	45.6 \pm 4.6	5.47 \pm 0.53	5.10 \pm 0.92
	hsa-miR-99a	43.3 \pm 2.9	5.70 \pm 1.10	3.67 \pm 0.16
	hsa-miR-518f	42.2 \pm 1.7	4.56 \pm 0.25	4.05 \pm 0.37
	hsa-miR-885-5p	39.1 \pm 1.7	5.98 \pm 0.55	4.12 \pm 0.12
	hsa-miR-302d*	40.8 \pm 3.9	4.86 \pm 0.52	5.33 \pm 0.22
	cel-miR-67	6.6 \pm 1.1	1.18 \pm 0.15	0.22 \pm 0.02
	hsa-miR-199a-3p	38.1 \pm 1.4	8.46 \pm 0.79	2.99 \pm 0.48
Mouse CMs	hsa-miR-33b*	29.9 \pm 3.2	5.12 \pm 0.81	1.40 \pm 0.18
	hsa-miR-590-3p	25.2 \pm 1.2	5.14 \pm 1.11	2.10 \pm 0.24
	hsa-miR-1244	20.7 \pm 1.4	5.76 \pm 0.21	1.35 \pm 0.04
	hsa-miR-23b	21.1 \pm 1.3	5.33 \pm 0.52	2.56 \pm 0.23
	hsa-miR-152	24.1 \pm 1.7	5.83 \pm 0.87	1.29 \pm 0.33
	hsa-miR-19b-2*	28.5 \pm 0.9	4.48 \pm 0.48	2.41 \pm 0.57
	hsa-miR-372	21.0 \pm 1.0	4.80 \pm 0.12	1.53 \pm 0.20
	hsa-miR-130a*	22.4 \pm 1.5	4.63 \pm 0.20	1.55 \pm 0.25
	hsa-miR-885-5p	17.3 \pm 0.6	4.75 \pm 0.32	1.52 \pm 0.27

Percentage of proliferating (Ki-67⁺ and EdU⁺), H3S10ph⁺ and midbodies⁺ rat and mouse CMs (α -actinin⁺) transfected with the 10 miRNAs that increase proliferation of rat and mouse CMs more efficiently. Results for the control cel-miR-67 are shown for comparison. Data are shown as mean values \pm s.e.m.

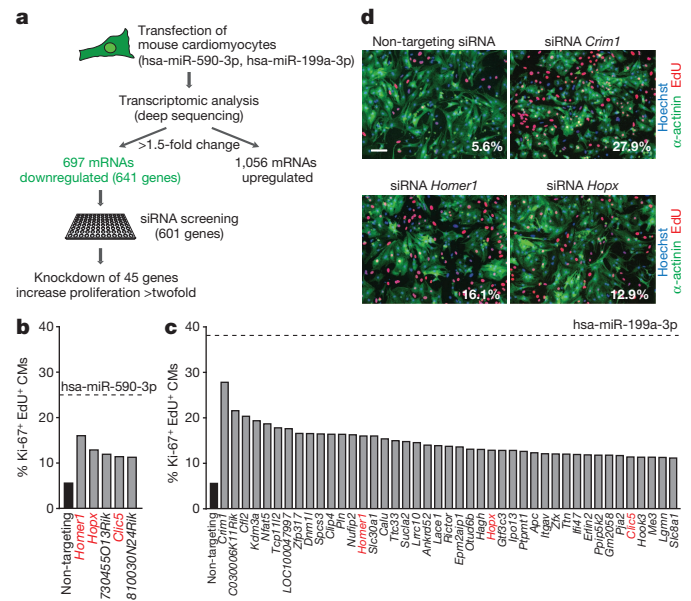
† Data are reported to the total number of CMs.

miRNA effect depends on multiple targets

To identify the relevant targets of the selected miRNAs, we assessed global transcriptome changes by deep-sequencing total neonatal mouse CM RNA after transfection with hsa-miR-590-3p or hsa-miR-199a-3p, which are the two miRNAs that were the most effective in rat and mouse CMs, respectively.

This analysis identified 1,056 upregulated transcripts (283 for hsa-miR-590-3p and 773 for hsa-miR-199a-3p; reads per kilobase of exon model per million mapped reads (RPKM) > 5 and 1.5-fold upregulation) and 697 downregulated transcripts (95 for hsa-miR-590-3p and 602 for hsa-miR-199a-3p; RPKM > 5 and 1.5-fold downregulation, including 35 mRNAs downregulated by both miRNAs; these corresponded to 641 different genes; Supplementary Table 2 and Fig. 3a). Functional analysis of the upregulated transcripts revealed enrichment for genes belonging to the 'cell cycle', 'cellular growth and proliferation' and 'DNA replication, recombination and repair' categories (Supplementary Fig. 11a). Analysis of the 641 downregulated genes showed enrichment for genes belonging to the 'skeletal and muscular system development and function' and 'cellular assembly and organization' categories (Supplementary Fig. 11b).

To identify, among the downregulated genes, those relevant in controlling CM proliferation, we individually knocked down 601 of these genes for which short interfering RNAs were available (Fig. 3a), under conditions similar to those of the initial screening. Forty-five of these siRNAs increased proliferation at least twofold (Supplementary Table 3). Of these, 5 siRNAs targeted transcripts downregulated by hsa-miR-590-3p (Fig. 3b) and 43 by hsa-miR-199a-3p (Fig. 3c); 3 siRNAs targeted transcripts downregulated by both miRNAs (red in Fig. 3b, c); Fig. 3d shows representative images from these experiments. The transcripts downregulated by both miRNAs were *Homer1*, involved in the regulation of calcium signalling in cardiac cells³⁰, *Hopx*, a regulator suppressing embryonic CM proliferation³¹, and *Clic5*. Luciferase 3'-untranslated region (UTR) reporter assays showed that *Homer1* and *Hopx* are directly regulated by hsa-miR-590-3p and



miRNAs in increasing CM proliferation probably ensues as a cumulative effect on multiple, cellular mRNA targets.

miRNAs increase CM proliferation *in vivo*

We went on to assess the *in vivo* effect of hsa-miR-590-3p and hsa-miR-199a-3p. The synthetic miRNAs, complexed with a lipid transfection reagent, were injected directly into the heart of neonatal rats. One day later, the animals were administered EdU intraperitoneally (Fig. 4a). After three additional days, the left ventricle walls of the hearts injected with the two miRNAs appeared markedly thicker than those that received cel-miR-67 (Fig. 4b). Analysis of EdU incorporation revealed a marked increase in the number of EdU⁺ cells in these hearts (Fig. 4b, c and Supplementary Fig. 13a). Confocal microscopy analysis indicated that the majority of the proliferating cells were indeed CMs (Fig. 4d). Consistent with the lack of induction of proliferation on cardiac fibroblasts *in vitro*, no increased fibrosis was observed in the injected hearts (Supplementary Fig. 13b).

To explore the long-term effects of the miRNAs in the neonatal heart, we exploited the high cardiac tropism and prolonged expression of viral vectors based on adeno-associated virus (AAV) serotype 9 upon systemic delivery^{32–34}. AAV9 vectors expressing hsa-miR-590 or hsa-miR-199a precursor miRNAs, or a control vector, were injected intraperitoneally in neonatal mice (Fig. 4e). At 12 days after injection, the hearts of these animals were morphologically normal, but significantly enlarged (Fig. 4f, g). No difference in collagen content was observed (data not shown). Similarly, there were no signs of CM hypertrophy (Fig. 4h and Supplementary Fig. 14).

The number of mitotic CMs, positive for H3S10ph, was significantly increased in the hearts of the animals injected with AAV9-miR-590 or AAV9-miR-199a (Fig. 4i and Supplementary Fig. 15). Confocal microscopy analysis of the proliferating cells revealed that most of these cells were differentiated CMs; however, the cells showed partial sarcomere disassembly, as would be expected in cells undergoing division³⁵ (Fig. 4j). Intracardiac delivery of these miRNAs in adult mice also resulted in increased cardiac cell proliferation (Supplementary Fig. 16a–c). AAV vector transduction determined a significant overexpression of the miRNAs both in neonatal and adult animals (Supplementary Fig. 17a–d).

Collectively, these results show that hsa-miR-590 and hsa-miR-199a significantly increase CM proliferation *in vivo*.

miRNAs preserve cardiac function after infarction

Next we assessed whether hsa-miR-590 and hsa-miR-199a might boost the normally ineffective myocardial repair that takes place after myocardial infarction. Adult female CD1 mice (8–12-weeks old) underwent permanent left anterior descending coronary artery ligation and were injected, in the peri-infarcted area, with AAV9 vectors expressing the two miRNAs or a control vector (Fig. 5a). Our previous experience indicates that this treatment results in efficient myocardial transduction and month-long transgene expression³⁶. As evaluated by echocardiography, the left ventricular ejection fraction (LVEF, Fig. 5b), fractional shortening (LVFS, Fig. 5c), end-systolic anterior wall thickness (LVAW thickness-systolic, Fig. 5d), as well as other parameters of cardiac function (Supplementary Fig. 18a–c), were significantly preserved over time in the infarcted mice injected with AAV9-miR-590 or AAV9-miR-199a.

Analysis of trichrome-stained heart cross-sections clearly showed that the infarct size was significantly reduced in mice treated with both miRNAs (Fig. 5e, f). In the animals treated with the two miRNAs, at 60 days after vector delivery and EdU treatment, a significant number of EdU-positive nuclei was still detected in the infarct border zone (Fig. 5g). Confocal microscopy revealed that most of these nuclei belonged to mature CMs, well integrated within the myocardial structure (Fig. 5h), indicative of active proliferation soon after infarction and treatment. Of note, administration of hsa-miR-590 and hsa-miR-199a did not protect CMs from

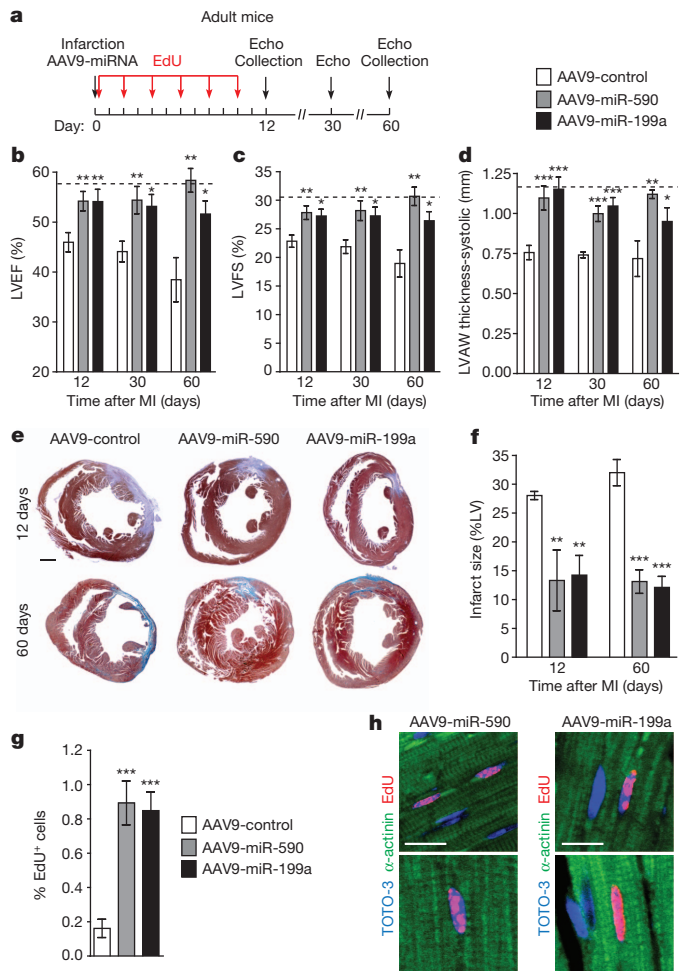


Figure 5 | miR-590-3p and miR-199a-3p induce marked cardiac regeneration after myocardial infarction. **a**, Schematic of the myocardial infarction (MI) experiments in adult mice. Echo, echocardiography. **b–d**, LVEF (**b**), LVFS (**c**) and LVAW thickness-s (**d**). Dashed line, non-infarcted animals injected with AAV9-control. *n* = 10–16 per group. **e**, Masson trichrome staining of heart cross sections. Scale bar, 1 mm. **f**, Infarct size. *n* = 6–10 per group. **g**, **h**, Percentage of cells positive for EdU (**g**) and confocal microscopy (**h**) at the border zone of the infarct, 60 days post-MI. Scale bar, 10 μ m. All panels, mean \pm s.e.m.; **P* < 0.05, ***P* < 0.01, ****P* < 0.001 relative to control.

cardiotoxic stimuli *in vitro* (Supplementary Fig. 19a), nor did it have a cardioprotective effect after myocardial infarction (Supplementary Fig. 19b).

Together, these data indicate that the expression of hsa-miR-590 and hsa-miR-199a after infarction exerts a marked beneficial effect in reducing infarct size and improving cardiac function, consistent with the effect of these miRNAs in actively stimulating CM proliferation.

Discussion

To the best of our knowledge, this is the first demonstration that CM proliferation can be stimulated by the exogenous administration of miRNAs and, most pertinently, that such treatment can restore cardiac mass and promote functional recovery after myocardial infarction in adult animals. The marked proliferative effect exerted by these miRNAs likely ensues as the sum of their effects on multiple, individual genes.

An important question prompted by our study concerns the identity of the cells upon which hsa-miR-590-3p and hsa-miR-199a-3p act once delivered into the heart *in vivo*. In the neonate, depending on the species, approximately 3–15% of the CMs are still in the cell cycle (see refs. 24, 25, 37 and this manuscript), and thus it might be envisaged that the miRNAs enhance the proliferative capacity of these cells.

Similarly, hsa-miR-590-3p and hsa-miR-199a-3p could act by increasing the normally limited CM proliferation that occurs in the adult heart in the infarct border zone. Intriguingly, however, both these miRNAs also induced re-entry into the cell cycle of adult CMs *ex vivo* and promoted DNA synthesis and the appearance of mitotic figures in CMs *in vivo*. These observations support the conclusion that these miRNAs promote proliferation of already differentiated CMs, a condition known to associate with cardiac regeneration in zebrafish^{38–40} and in the neonatal mouse⁴¹.

Regeneration of the cardiac mass seems important in a variety of pathological conditions associated with CM loss, leading from myocardial infarction to include heart failure and dilated cardiomyopathy^{42,43}. Thus, the findings that exogenously administered miRNAs have the potential to restore cardiac function to almost normal levels after myocardial infarction have evident translational value. These miRNAs, in addition, might be useful in expanding the proliferative potential of CMs derived from human cardiac stem cells, which have shown promise in recent clinical trials^{44,45}.

METHODS SUMMARY

miRNA library screening. Ventricular CMs from neonatal rats were transfected with a library of miRNA mimics (Dharmacon) using a standard reverse transfection protocol. EdU was added 52 h later; 72 h after plating cells were fixed and stained. Image acquisition was performed using an ImageXpress Micro high-content screening microscope (Molecular Devices). CMs were scored as proliferating only if positive for the two proliferation markers (Ki-67 and EdU).

Animal studies. Animals were injected with synthetic miRNAs complexed with a transfection reagent or with AAV9 vectors encoding the microRNAs. Proliferation of cardiac cells was analysed 4 and 12 days after injection. Myocardial infarction was induced in adult CD1 mice by permanent left anterior descending coronary artery ligation; cardiac function was evaluated by histological, morphological and echocardiographic analysis.

Full Methods and any associated references are available in the online version of the paper.

Received 29 March; accepted 31 October 2012.

Published online 5 December 2012.

- Ahuja, P., Sdek, P. & MacLellan, W. R. Cardiac myocyte cell cycle control in development, disease, and regeneration. *Physiol. Rev.* **87**, 521–544 (2007).
- van Amerongen, M. J. & Engel, F. B. Features of CM proliferation and its potential for cardiac regeneration. *J. Cell. Mol. Med.* **12**, 2233–2244 (2008).
- Bicknell, K. A., Coxon, C. H. & Brooks, G. Can the CM cell cycle be reprogrammed? *J. Mol. Cell. Cardiol.* **42**, 706–721 (2007).
- Bergmann, O. *et al.* Evidence for CM renewal in humans. *Science* **324**, 98–102 (2009).
- Kajstura, J. *et al.* Cardiomyogenesis in the adult human heart. *Circ. Res.* **107**, 305–315 (2010).
- Beltrami, A. P. *et al.* Evidence that human cardiac myocytes divide after myocardial infarction. *N. Engl. J. Med.* **344**, 1750–1757 (2001).
- Robledo, M. Myocardial regeneration in young rats. *Am. J. Pathol.* **32**, 1215–1239 (1956).
- Nag, A. C., Carey, T. R. & Cheng, M. DNA synthesis in rat heart cells after injury and the regeneration of myocardia. *Tissue Cell* **15**, 597–613 (1983).
- Kajstura, J. *et al.* Myocyte cellular hyperplasia and myocyte cellular hypertrophy contribute to chronic ventricular remodeling in coronary artery narrowing-induced cardiomyopathy in rats. *Circ. Res.* **74**, 383–400 (1994).
- Reiss, K., Kajstura, J., Capasso, J. M., Marino, T. A. & Anversa, P. Impairment of myocyte contractility following coronary artery narrowing is associated with activation of the myocyte IGF1 autocrine system, enhanced expression of late growth related genes, DNA synthesis, and myocyte nuclear mitotic division in rats. *Exp. Cell Res.* **207**, 348–360 (1993).
- Bartel, D. P. miRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).
- Pasquinelli, A. E. miRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nature Rev. Genet.* **13**, 271–282 (2012).
- Chen, J. F. *et al.* Targeted deletion of Dicer in the heart leads to dilated cardiomyopathy and heart failure. *Proc. Natl Acad. Sci. USA* **105**, 2111–2116 (2008).
- Rao, P. K. *et al.* Loss of cardiac miRNA-mediated regulation leads to dilated cardiomyopathy and heart failure. *Circ. Res.* **105**, 585–594 (2009).
- Ikeda, S. *et al.* Altered miRNA expression in human heart disease. *Physiol. Genomics* **31**, 367–373 (2007).
- Matkovich, S. J. *et al.* Reciprocal regulation of myocardial miRNAs and messenger RNA in human cardiomyopathy and reversal of the miRNA signature by biomechanical support. *Circulation* **119**, 1263–1271 (2009).
- Thum, T. *et al.* MiRNAs in the human heart: a clue to fetal gene reprogramming in heart failure. *Circulation* **116**, 258–267 (2007).
- van Rooij, E. *et al.* A signature pattern of stress-responsive miRNAs that can evoke cardiac hypertrophy and heart failure. *Proc. Natl Acad. Sci. USA* **103**, 18255–18260 (2006).
- Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific miRNA that targets *Hand2* during cardiogenesis. *Nature* **436**, 214–220 (2005).
- Zhao, Y. *et al.* Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1–2. *Cell* **129**, 303–317 (2007).
- Liu, N. *et al.* miRNA-133a regulates CM proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev.* **22**, 3242–3254 (2008).
- Porrello, E. R. *et al.* MiR-15 family regulates postnatal mitotic arrest of CMs. *Circ. Res.* **109**, 670–679 (2011).
- Salic, A. & Mitchison, T. J. A chemical method for fast and sensitive detection of DNA synthesis *in vivo*. *Proc. Natl Acad. Sci. USA* **105**, 2415–2420 (2008).
- Soonpaa, M. H., Kim, K. K., Pajak, L., Franklin, M. & Field, L. J. CM DNA synthesis and binucleation during murine development. *Am. J. Physiol.* **271**, H2183–H2189 (1996).
- Li, F., Wang, X., Capasso, J. M. & Gerdes, A. M. Rapid transition of cardiac myocytes from hyperplasia to hypertrophy during postnatal development. *J. Mol. Cell. Cardiol.* **28**, 1737–1746 (1996).
- Kubin, T. *et al.* Oncostatin M is a major mediator of CM dedifferentiation and remodeling. *Cell Stem Cell* **9**, 420–432 (2011).
- Fredji, S., Bescond, J., Louault, C. & Potreau, D. Interactions between cardiac cells enhance CM hypertrophy and increase fibroblast proliferation. *J. Cell. Physiol.* **202**, 891–899 (2005).
- Zhang, Y. *et al.* Dedifferentiation and proliferation of mammalian CMs. *PLoS ONE* **5**, e12559 (2010).
- Bird, S. D. *et al.* The human adult CM phenotype. *Cardiovasc. Res.* **58**, 423–434 (2003).
- Pouliquin, P. & Dulhunty, A. F. Homer and the ryanodine receptor. *Eur. Biophys. J.* **39**, 91–102 (2009).
- Trivedi, C. M. *et al.* Hopx and Hdac2 interact to modulate Gata4 acetylation and embryonic cardiac myocyte proliferation. *Dev. Cell* **19**, 450–459 (2010).
- Inagaki, K. *et al.* Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol. Ther.* **14**, 45–53 (2006).
- Katara, R. *et al.* Intravenous gene therapy with PIM-1 via a cardiotropic viral vector halts the progression of diabetic cardiomyopathy through promotion of pro-survival signaling. *Circ. Res.* **108**, 1238–1251 (2011).
- Secchiuro, P. *et al.* Systemic tumor necrosis factor-related apoptosis-inducing ligand delivery shows antiatherosclerotic activity in apolipoprotein E-null diabetic mice. *Circulation* **114**, 1522–1530 (2006).
- Ahuja, P., Perriard, E., Perriard, J. C. & Ehler, E. Sequential myofibrillar breakdown accompanies mitotic division of mammalian CMs. *J. Cell Sci.* **117**, 3295–3306 (2004).
- Zentilin, L. *et al.* CM VEGFR-1 activation by VEGF-B induces compensatory hypertrophy and preserves cardiac function after myocardial infarction. *FASEB J.* **24**, 1467–1478 (2010).
- Collesi, C., Zentilin, L., Sinagra, G. & Giacca, M. Notch1 signaling stimulates proliferation of immature CMs. *J. Cell Biol.* **183**, 117–128 (2008).
- Jopling, C. *et al.* Zebrafish heart regeneration occurs by CM dedifferentiation and proliferation. *Nature* **464**, 606–609 (2010).
- Kikuchi, K. *et al.* Primary contribution to zebrafish heart regeneration by *gata4*⁺ CMs. *Nature* **464**, 601–605 (2010).
- Poss, K. D., Wilson, L. G. & Keating, M. T. Heart regeneration in zebrafish. *Science* **298**, 2188–2190 (2002).
- Porrello, E. R. *et al.* Transient regenerative potential of the neonatal mouse heart. *Science* **331**, 1078–1080 (2011).
- Dorn, G. W. II. Apoptotic and non-apoptotic programmed CM death in ventricular remodeling. *Cardiovasc. Res.* **81**, 465–473 (2009).
- Nishida, K. & Otsu, K. Cell death in heart failure. *Circ. J.* **72** (Suppl. A) A17–A21 (2008).
- Bolli, R. *et al.* Cardiac stem cells in patients with ischaemic cardiomyopathy (SCIPIO): initial results of a randomised phase 1 trial. *Lancet* **378**, 1847–1857 (2011).
- Makkar, R. R. *et al.* Intracoronary cardiosphere-derived cells for heart regeneration after myocardial infarction (CADUCEUS): a prospective, randomised phase 1 trial. *Lancet* **379**, 895–904 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors are grateful to M. Dapas and M. Zotti for technical support in AAV production, to M. Sturmege for help in animal experimentation and to S. Kerbavcic for editorial assistance. A.E. is recipient of a FEBS Long Term Fellowship. This work was supported by Advanced Grant 250124 from the European Research Council (ERC) to M.G. and from Project CTC from the Fondazione CRTrieste, Trieste, Italy.

Author Contributions A.E., M.M. and M.G. designed the study. A.E., M.M., S.Z. and M.D.F. performed the experiments and analysed the data. G.S. provided clinical consultancy for the animal study. L.Z. was responsible for AAV production. A.E., M.M. and M.G. wrote the manuscript.

Author Information The cardiomyocyte miRNA expression microarray and transcriptomic data are deposited at GEO, under accession numbers GSE41537 and GSE41538, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G. (giacca@icgeb.org).

METHODS

Isolation of ventricular CMs from neonatal and 7-days-old animals. Wistar rats and CD1 mice were purchased from Charles River Laboratories Italia Srl. Animal care and treatments were conducted in conformity with institutional guidelines in compliance with national and international laws and policies (EEC Council Directive 86/609, OJL 358, 12 December 1987).

Ventricular CMs from neonatal rats or mice were isolated as described previously³⁷, with minor modifications. In brief, ventricles from neonatal rats or mice (post-natal day 0) were separated from the atria, cut into pieces and then dissociated in calcium and bicarbonate-free Hanks with HEPES (CBFHH) buffer containing 1.75 mg ml⁻¹ trypsin (BD Difco) and 10 µg ml⁻¹ DNase II (Sigma), under constant stirring. Digestion was performed at room temperature in eight to ten 10-min steps, collecting the supernatant to fetal bovine serum (FBS, Life Technologies) after each step. The collected supernatant was centrifuged to separate the cells, which were then resuspended in Dulbecco's modified Eagle medium 4.5 g l⁻¹ glucose (DMEM, Life Technologies) supplemented with 5% FBS, 20 µg ml⁻¹ vitamin B12 (Sigma) and with 100 U ml⁻¹ of penicillin and 100 µg ml⁻¹ of streptomycin (Sigma). The collected cells were passed through a cell strainer (40 µm, BD Falcon) and then seeded onto uncoated 100-mm plastic dishes for 2 h at 37 °C in 5% CO₂ and humidified atmosphere. The supernatant, composed mostly of CMs, was then collected and pelleted. Cells were resuspended in antibiotic-free media, counted and plated at the appropriate density; cultures of neonatal rat or mouse ventricular CMs prepared using this procedure yielded consistently a purity of >90%.

Ventricular CMs from 7-days-old Wistar rats were isolated essentially as described for neonatal rats, except that digestions were performed at 37 °C in CBFHH buffer containing 0.25 mg ml⁻¹ of pancreatin (Sigma), 0.125 mg ml⁻¹ collagenase type II (Worthington Biochemical) and 10 µg ml⁻¹ DNase II (Sigma), under constant stirring. Cells were washed thoroughly 24 h after seeding, and transfected with the miRNAs, as described below.

Isolation of adult rat ventricular CMs. Ventricular CMs were isolated from Langendorff-perfused hearts of adult female Wistar rats (2-months-old) as previously described⁴⁶, with minor modifications. Briefly, hearts were extracted and perfused retrogradely with calcium-free Krebs–Henseleit bicarbonate (KHB) buffer. Hearts were then perfused with KHB buffer containing 1 mg ml⁻¹ Liberase (Roche) for 10 min. Following removal of the atria and great vessels, the hearts were minced in KHB buffer and the cell mixture was filtered through a cell strainer (100 µm, BD Falcon). The cells were then pelleted by centrifugation at 530g for 3 min at room temperature. The cell pellet was resuspended in a mixture DMEM 1.0 g l⁻¹ glucose (Life Technologies) and perfusion buffer (1:1) and the separation of CMs from other cell types was achieved by sedimentation on a 6% bovine serum albumin (BSA, Sigma) cushion for 15 min. The CM pellet was resuspended and plated in DMEM 1.0 g l⁻¹ glucose supplemented with 2 g l⁻¹ BSA, 2 mM L-carnitine (Sigma), 5 mM creatine (Sigma), 5 mM taurine (Sigma), 1 mM 2,3-butanedione monoxime (BDM; Sigma) and with 100 U ml⁻¹ of penicillin and 100 µg ml⁻¹ of streptomycin. Cells were plated on 24-well plates coated with laminin (Sigma), and kept at 37 °C in 5% CO₂ and humidified atmosphere. The medium was exchanged 24 h later to DMEM 4.5 g l⁻¹ glucose supplemented with 5% FBS, 20 µg ml⁻¹ vitamin B12 and the cells were transfected as described below.

miRNA transfection. The miRNA mimics library (miRIDIAN miRNA mimics) corresponding to all the human mature miRNAs (988 miRNAs, 875 unique sequences, miRBase 13.0), was obtained from Dharmacon, Thermo Scientific. Sequences of all miRNAs and miRBase accession numbers are included in Supplementary Table 1. For the miRNA screening in rat CMs, miRNAs were transferred robotically from stock library plates to Primaria 96-well plates (BD Falcon) leaving columns 1 and 12 empty for addition of controls (buffer, cel-miR-67 and rno-miR-1). MiRNAs were transfected into neonatal rat CMs using a standard reverse transfection protocol, at a final miRNA concentration of 25 nM. Briefly, the transfection reagent (Lipofectamine RNAiMAX, Life Technologies) was diluted in OPTI-MEM (Life Technologies) and added to the miRNAs arrayed on 96-well plates; 30 min later, 1.0×10^4 cells were seeded per well. Screening was performed in duplicate. Twenty-four hours after transfection, culture medium was replaced by fresh medium; 28 h later, that is, 52 h after plating, the culture medium was replaced with medium containing 5 µM 5-ethynyl-2'-deoxyuridine (EdU, Life Technologies) for 20 h. Cells were fixed at 72 h after plating and processed for immunofluorescence.

For the transfection experiments in mouse CMs, the selected miRNAs were 'cherry-picked' from miRNA library stock plates or transferred from individual tubes and re-arrayed onto collagen-coated black clear-bottom 96-well plates (PerkinElmer). Transfection was performed as described above, except that 1.5×10^4 cells were seeded per well and a final miRNA concentration of 50 nM was used. For RNA isolation, 9×10^5 cells were seeded in Primaria 60-mm dishes

(BD Falcon) and transfected with miRNA mimics at a final concentration of 50 nM.

The screenings were performed at the ICGEB High-Throughput Screening Facility (<http://www.icgeb.org/high-throughput-screening.html>).

Transfection of rat CMs isolated from 7-days-old and adult animals was performed essentially as described above for neonatal CMs, except that transfections were performed 24 h after cell seeding using a forward transfection protocol at a final miRNA concentration of 25 nM and 50 nM, respectively. For the 7-days-old CMs the medium was replaced with medium containing 5 µM EdU 52 h after transfection, and cells were fixed 20 h later. For the adult CMs, the medium was replaced by medium containing 5 µM EdU 48 h after transfection and every 24 h thereafter until the cells were fixed, 6 days after transfection.

siRNA transfection. For the siRNA transfection, selected siRNAs were 'cherry-picked' from the mouse genome-wide siRNA library stock plates (siGENOME SMARTpools, Dharmacon, Thermo Scientific; gene IDs and gene symbols are included in Supplementary Table 3) and re-arrayed onto collagen-coated black clear-bottom 384-well plates (PerkinElmer). siRNA transfection was performed at a final concentration of 50 nM, as described previously for the miRNA screening, except that 7.5×10^3 cells were seeded per well. Cells were fixed at 72 h after plating and processed for immunofluorescence.

Hypertrophy studies. For a detailed analysis of the effect of miRNAs on CM size, 1.5×10^4 cells were seeded per well on Primaria 96-well plates (for microscopy) or 6×10^5 cells were seeded in Primaria 60-mm dishes (for quantitative PCR); 24 h after cell seeding, the medium was exchanged to low-serum culture medium, specifically DMEM 4.5 g l⁻¹ glucose supplemented with 0.1% FBS, and transfected as described above. Cells were fixed 72 h after transfection and processed for immunofluorescence or qPCR, as described below.

Luciferase 3'-UTR reporter assays. Constructs bearing the murine 3'-UTR of *Homer1*, *Hopx* and *Clic5* were obtained by gene synthesis from BlueHeron Biotech and subcloned into psiCHECK2 vector (Promega). HeLa cells were transfected with cel-miR-67, hsa-miR-199a-3p or hsa-miR-590-3p at a final concentration of 50 nM in 96-well plates, through a standard reverse transfection protocol similar to that described above for CMs. Twenty-four hours after miRNA transfection, cells were transfected with 100 ng per well of the reporter constructs or psiCHECK2 vector (control) using FuGENE HD transfection reagent (Promega). Firefly and *Renilla* luciferase activities were measured 48 h after plasmid transfection using the Dual Luciferase Reporter Assay System (Promega), according to the manufacturer's instructions.

Intracardiac injection of miRNA complexes in neonatal rats. Neonatal Wistar rats (post-natal day 0) were anesthetized by cooling on an ice bed for 5 min. Lateral thoracotomy at the fourth intercostal space was performed by dissection of the intercostal muscles following skin incision. Following intracardiac injection of the miRNA complexes using an insulin syringe with incorporated 30-gauge needle (Roche), the animals were removed from the ice bed and thoracic wall and skin incisions were sutured. Rats were then placed under a heat lamp and warmed for several minutes until recovery. The miRNA complexes were prepared by mixing the miRNA (approximately 2.8 µg of cel-miR-67, hsa-miR-590-3p or hsa-miR-199a-3p, Dharmacon, Thermo Scientific) with Lipofectamine RNAiMAX transfection reagent for 30 min at room temperature. The total volume of the mix injected per heart was 30 µl. EdU was administered intraperitoneally (350 µg per animal) 2 days after miRNA injection. The hearts of the injected rats ($n = 8$ per group) were collected 4 days after miRNA injection.

Production and purification of recombinant AAV vectors. The precursors of hsa-miR-199a and hsa-miR-590 plus upstream and downstream flanking sequences (total approximating 300 base pairs) were amplified from human genomic DNA isolated from HeLa cells, using the QIAamp DNA mini kit (Qiagen), according to the manufacturer's instructions. The primers used to amplify the precursor sequences were: hsa-miR-199a, forward 5'-CGGAATTC GCTCCGCTCCCCACTCTTTAG-3', reverse 5'-GCTCTAGAATGGGGTG GGGATGGCAGACTGA; hsa-miR-590, forward 5'-CGGAATTC CCCCCCAT CAAAACAAAACCGGTGT-3', reverse 5'-GCTCTAGAACTCCATCGAGTG AGCCACA-3'. The amplified sequences were cloned into the pZac2.1 vector (Gene Therapy Program, Penn Vector core, University of Pennsylvania, USA), which was used to produce recombinant AAV vectors.

Recombinant AAV vectors were prepared in the AAV Vector Unit at ICGEB Trieste, as described previously⁴⁷. Briefly, AAV vectors of serotype 9 were generated in HEK293T cells, using a triple-plasmid co-transfection for packaging. Viral stocks were obtained by CsCl₂ gradient centrifugation. Titration of AAV viral particles was performed by real-time PCR quantification of the number of viral genomes, as described previously; the viral preparations had titres between 1×10^{13} and 3×10^{13} viral genome particles per ml.

Injection of AAV vectors in neonatal and adult mice. In the experiments using AAV vectors, neonatal CD1 mice (post-natal day 1) were intraperitoneally

injected with AAV9-LacZ (AAV9-control), AAV9-miR-590 or AAV9-miR-199a at a dose of 1×10^{11} viral genome particles per animal, using an insulin syringe with incorporated 30-gauge needle. The hearts of the injected mice ($n = 9$ per group) were collected 12 days after AAV injection.

Adult mice intracardiac injection of AAV vectors (AAV9-control, AAV9-miR-590 or AAV9-miR-199a), at a dose of 1×10^{11} viral genome particles per animal, was performed as described below for animals that underwent myocardial infarction. The hearts of the injected animals ($n = 3$ per group) were collected 12 days after AAV injection.

Myocardial infarction. Myocardial infarction was produced in adult female CD1 mice (8–12 weeks old), by permanent left anterior descending (LAD) coronary artery ligation. Briefly, mice were anesthetized with an intraperitoneally injection of ketamine and xylazine, endotracheally intubated and placed on a rodent ventilator. Body temperature was maintained at 37°C on a heating pad. The beating heart was accessed via a left thoracotomy. After removing the pericardium, a descending branch of the LAD coronary artery was visualized with a stereomicroscope (Leica) and occluded with a nylon suture. Ligation was confirmed by the whitening of a region of the left ventricle, immediately post-ligation. Recombinant AAV vectors, at a dose of 1×10^{11} viral genome particles per animal, were injected immediately after LAD ligation into the myocardium bordering the infarct zone (single injection), using an insulin syringe with incorporated 30-gauge needle. Three groups of animals were studied, receiving AAV9-control, AAV9-miR-590 or AAV9-miR-199a. The chest was closed, and the animals moved to a prone position until the occurrence of spontaneous breathing. EdU was administered intraperitoneally (500 μg per animal) every 2 days, for a period of ten days. Echocardiography analysis was performed at days 12, 30 and 60 after infarction, as described below, and hearts were collected at 12 ($n = 6$ animals per group) and 60 ($n = 10$ animals per group) days after infarction.

Echocardiography analysis. To evaluate left ventricular function and dimensions, transthoracic two-dimensional echocardiography was performed on mice sedated with 5% isoflurane at 12, 30 and 60 days after myocardial infarction, using a Visual Sonics Vevo 770 Ultrasound (Visual Sonics) equipped with a 30-MHz linear array transducer. M-mode tracings in parasternal short axis view were used to measure left ventricular anterior and posterior wall thickness and left ventricular internal diameter at end-systole and end-diastole, which were used to calculate left ventricular fractional shortening and ejection fraction.

Heart collection and histological analysis. At the end of the studies, animals were anaesthetized with 5% isoflurane and then killed by injection of 10% KCl, to stop the heart at diastole. The heart was excised, briefly washed in PBS, weighted, fixed in 10% formalin at room temperature, embedded in paraffin and further processed for histology or immunofluorescence. Haematoxylin–eosin and Masson's trichrome staining were performed according to standard procedures, and analysed for regular morphology and extent of fibrosis. Infarct size was measured as the percentage of the total left ventricular area showing fibrosis.

Immunofluorescence. Cells were fixed with 4% paraformaldehyde for 15 min, permeabilized with 0.5% Triton X-100 in phosphate buffered saline (PBS) solution for 10 min, followed by 30 min blocking in 1% BSA (Roche). Cells were then stained overnight at 4°C with the following primary antibodies diluted in blocking solution: mouse monoclonal antibody against sarcomeric α -actinin (Abcam), rabbit antibodies against Ki-67 (Monosan), histone H3 phosphorylated at serine 10 (Millipore), Aurora B kinase (Sigma) or atrial natriuretic peptide (ANP, Millipore). Cells were washed with PBS and incubated for 2 h with the respective secondary antibodies conjugated to Alexa Fluor-488, -555 or -647 (Life Technologies). When indicated, cells were further processed using the Click-IT EdU 555 Imaging kit to reveal EdU incorporation, according to the manufacturer's instructions, and stained with Hoechst 33342 (Life Technologies).

Sections of rat or mouse hearts were deparaffinized, rehydrated and then underwent antigen retrieval by boiling in sodium citrate solution for 20 min. Slides were processed for immunofluorescence as described for cultured CMs. Nuclei were identified by counter-staining sections with TOTO-3 (Life Technologies) and slides were then mounted in Vectashield with DAPI (Vector Labs).

For wheat germ agglutinin (WGA) staining, heart sections were deparaffinized, rehydrated and then incubated for 1 h at room temperature with WGA conjugated to Alexa Fluor-594 (50 $\mu\text{g ml}^{-1}$, Life Technologies) in PBS. Slides were then rinsed in PBS and mounted in Vectashield.

Image acquisition and analysis. For the screening experiments, image acquisition was performed using an ImageXpress Micro automated high-content screening fluorescence microscope at $\times 20$ magnification; a total of 16 images were acquired per wavelength, well and replicate, corresponding to approximately 2,500 cells analysed per condition and replicate. Image analysis was performed using the 'Multi-Wavelength Cell Scoring' application module implemented in MetaXpress software (Molecular Devices). Cells were scored as proliferating only if positive for the two proliferation markers (Ki-67 and EdU). Quantification of

cells positive for histone H3 phosphorylated at serine 10, ANP and terminal deoxynucleotidyl transferase dUTP nick end labelling (TUNEL) was performed using the same methodology. Quantification of cells presenting Aurora B in midbodies was performed by manual inspection and counting of 25 images acquired per well, at a $\times 20$ magnification. In all quantifications, CMs were distinguished from other cells present in the primary cultures (for example, fibroblasts and endothelial cells) by their positivity for sarcomeric α -actinin.

Sections of rat or mouse hearts were imaged using an ImageXpress Micro automated high-content screening fluorescence microscope or a laser scanning confocal 510 META microscope (Carl Zeiss MicroImaging). Fluorescent images of whole hearts were obtained by stitching of individual images, using MetaXpress software.

RNA isolation and quantitative real-time PCR. Total RNA, including small RNA fraction, from isolated CMs or dissected ventricular heart tissue samples was extracted using miRNeasy Mini Kit (Qiagen), according to the manufacturer's instructions.

For quantification of gene expression in isolated CMs, total RNA was reverse transcribed using hexameric random primers followed by qRT-PCR using pre-designed TaqMan assays (Applied Biosystems) for genes related to cardiac differentiation (*Myh6*, *Myh7*, *Nkx2.5*, *Dab2*, *destrin* and *Runx1*), cell cycle (*Ccne1*, *Ccna2* and *Ccnb1*) and hypertrophy (*ANP*, also known as *Nppa*), according to the manufacturer's instructions. The housekeeping gene *Gapdh* was used for normalization.

For miRNA quantification in samples of neonatal and adult mice at 12 days after intraperitoneal or intracardiac injection, respectively, RNA was transcribed using miRCURY LNA Universal cDNA synthesis kit (Exiqon) followed by qRT-PCR using pre-designed miRCURY LNA PCR primer sets (Exiqon) and miRCURY LNA SYBR Green master mix, according to the manufacturer's instructions. Expression of selected miRNAs was normalized to expression levels of 5S rRNA.

Fold changes were determined using the $2^{-\Delta\Delta C_t}$ method⁴⁸.

miRNA microarray. The miRNA profiling was performed by Exiqon Services, using a sixth generation miRCURY LNA miRNA Array (Exiqon), which contains capture probes targeting all rat miRNAs annotated in miRBase 16.0. Briefly, total RNA from neonatal or adult CMs, isolated as described above, was used for the hybridizations; the quality of the total RNA was verified by an Agilent 2100 Bioanalyzer profile. 750 ng total RNA from each sample was labelled with Hy3 fluorescent label, using the miRCURY LNA miRNA Hi-Power Labelling Kit, Hy3/Hy5 (Exiqon). The Hy3-labelled samples and a Hy5-labelled reference RNA sample (a mix of all the samples used for the study) were mixed pair-wise and hybridized to the array. The hybridization was performed using a Tecan HS 4800 hybridization station (Tecan). The miRCURY LNA array slides were scanned using the Agilent G2565BA Microarray Scanner System (Agilent Technologies) and the image analysis was carried out using the ImaGene 9.0 software (BioDiscovery). The quantified signals were background corrected (Normexp with offset value 10) and normalized using the global Lowess (locally weighted scatterplot smoothing) regression algorithm. Two-way hierarchical clustering was performed using Genesis software⁴⁹.

Transcriptomic analysis. For the transcriptomic analysis, mouse neonatal CMs were mock-transfected or transfected with cel-miR-67, hsa-miR-590-3p or hsa-miR-199a-3p as described above, except that 9.0×10^5 cells were seeded onto Primaria 60-mm dishes. Total RNA was extracted 72 h after transfection, using TRIzol (Life Technologies).

Deep-sequencing was performed by IGA Technology Services. Samples were processed using TruSeq RNA-seq sample prep kit from Illumina. Briefly, the poly-A containing mRNA molecules were purified using poly-T oligo-attached magnetic beads, fragmented into small pieces using divalent cations under elevated temperature, cDNA was synthesized by reverse transcription and standard blunt-ending plus added 'A' was performed. Then, Illumina TruSeq adapters with indexes were ligated to the ends of the cDNA fragments. After ligation and separation of non-ligated adapters, samples were amplified by PCR to selectively enrich those cDNA fragments in the library having adaptor molecules at both ends. A pool of 4 samples was loaded on one lane of an Illumina flowcell and clusters created by Illumina cBot. The clusters were sequenced at ultra-high-throughput on the Illumina HiSeq2000. One lane in 4-plex was run obtaining 35–40 millions of single-reads per sample, 50-bp long.

CLC-Bio Genomics Workbench software (CLC Bio) was used to calculate gene expression levels, as described previously⁵⁰. Transcript expression levels are expressed in terms of RPKM (reads per kilobase of exon model per million mapped reads). Transcripts with a RPKM > 5 in the mock-transfected cells were considered for further analysis. Fold-enrichment scores were obtained for each transcript by expressing the RPKM of a given transcript in CMs transfected with the selected miRNAs relative to mock-transfected cells. Only transcripts with a

variation of less than 20% between mock-transfected and cel-miR-67 were considered for the enrichment analysis. A fold enrichment cut-off of 1.5 was set.

Functional gene analysis. Functional analysis was conducted using Ingenuity Pathway Analysis Software (IPA, Ingenuity Systems) for the genes up or down-regulated in CMs transfected with hsa-miR-590-3p or hsa-miR-199a-3p by more than 1.5-fold compared to controls (1,056 and 697 transcripts, respectively). Data sets were uploaded into IPA and analysed for functional enrichment in terms of 'Molecular and Cellular Functions' and 'Physiological System Development and Function', based on the information in the Ingenuity Knowledge Base. Enrichment *P* values were calculated by IPA using the Fisher's exact test.

TUNEL assay. For the TUNEL analysis *in vitro*, 48 h after CM transfection with cel-miR-67, hsa-miR-590-3p or hsa-miR-199a-3p as described above, cells were mock treated or treated with 1 μ M doxorubicin or 1 μ M epirubicin for 20 h. For analysis of TUNEL *in vivo*, heart sections of animals collected 12 days after myocardial infarction and intracardiac AAV injection. TUNEL assay was performed using the Click-IT TUNEL Alexa Fluor-594 Imaging Assay (Life Technologies), according to the manufacturer's instructions.

Statistical analysis. All data are presented as mean \pm standard error of the mean (s.e.m.). Statistical analysis was carried out using Prism Software (GraphPad). For statistical comparison of two groups, unpaired, two-tailed Student's *t*-test was used; for the comparison of three or more groups, one-way ANOVA followed by Tukey's post-hoc test was used. A value of *P* < 0.05 was considered significant.

46. Xiao, L. *et al.* MEK1/2-ERK1/2 mediates α 1-adrenergic receptor-stimulated hypertrophy in adult rat ventricular myocytes. *J. Mol. Cell. Cardiol.* **33**, 779–787 (2001).
47. Arsic, N. *et al.* Induction of functional neovascularization by combined VEGF and angiopoietin-1 gene transfer using AAV vectors. *Mol. Ther.* **7**, 450–459 (2003).
48. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* **25**, 402–408 (2001).
49. Sturn, A., Quackenbush, J. & Trajanoski, Z. Genesis: cluster analysis of microarray data. *Bioinformatics* **18**, 207–208 (2002).
50. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5**, 621–628 (2008).

FMRP targets distinct mRNA sequence elements to regulate protein expression

Manuel Ascano Jr¹, Neelanjana Mukherjee^{2†}, Pradeep Bandaru¹, Jason B. Miller¹, Jeffrey D. Nusbaum¹, David L. Corcoran², Christine Langlois³, Mathias Munschauer¹, Scott Dewell⁴, Markus Hafner¹, Zev Williams^{1,3}, Uwe Ohler^{2†} & Thomas Tuschl¹

Fragile X syndrome (FXS) is a multi-organ disease that leads to mental retardation, macro-orchidism in males and premature ovarian insufficiency in female carriers. FXS is also a prominent monogenic disease associated with autism spectrum disorders (ASDs). FXS is typically caused by the loss of fragile X mental retardation 1 (*FMR1*) expression, which codes for the RNA-binding protein FMRP. Here we report the discovery of distinct RNA-recognition elements that correspond to the two independent RNA-binding domains of FMRP, in addition to the binding sites within the messenger RNA targets for wild-type and I304N mutant FMRP isoforms and the FMRP paralogues FXR1P and FXR2P (also known as FXR1 and FXR2). RNA-recognition-element frequency, ratio and distribution determine target mRNA association with FMRP. Among highly enriched targets, we identify many genes involved in ASD and show that FMRP affects their protein levels in human cell culture, mouse ovaries and human brain. Notably, we discovered that these targets are also dysregulated in *Fmr1*^{-/-} mouse ovaries showing signs of premature follicular overdevelopment. These results indicate that FMRP targets share signalling pathways across different cellular contexts. As the importance of signalling pathways in both FXS and ASD is becoming increasingly apparent, our results provide a ranked list of genes as basis for the pursuit of new therapeutic targets for these neurological disorders.

Most clinical cases of FXS are a result of a hyper-expansion and methylation of CGG repeats within the promoter of *FMR1*, leading to a loss of its expression^{1–3}. The FMR1 RNA-binding protein family has three members, FMRP, FXR1P and FXR2P, which possess two centrally located KH domains and a carboxy-terminal arginine-glycine-rich region implicated in mRNA binding^{4–7}. *FMR1* codes for multiple protein isoforms, but is predominantly expressed as a 69 kDa protein (isoform 7)^{8,9}. Isoform 1 and six other alternative splice variants include exon 12, with isoform 1 coding for the full-length protein (71 kDa). Exon 12 insertion lengthens the second KH (KH2) RNA-binding domain, possibly influencing FMRP RNA-binding specificity or affinity. The I304N mutation, first described in a FXS patient, is also located in the KH2 domain and is reported to attenuate association with RNA and polysomes^{10–12}. FMR1-family proteins are implicated in various RNA processes including RNA subcellular localization by facilitating nucleo-cytoplasmic shuttling¹³ and association with motor proteins^{14–16}, and are also suggested to mediate translational regulation^{12,17}. Given the critical role of *FMR1* in human cognition and premature ovarian insufficiency^{18,19}, there have been intensive efforts towards the identification of the RNA targets of FMRP, with the view that their discovery would shed light on the array of related disorders and provide options for molecular therapy^{18–26}. No precise RNA-recognition element (RRE) has been defined and very few bona fide mRNA targets are confirmed²⁷.

RNA target sites of the FMR1 protein family

To identify the binding sites of FMR1-family proteins (Fig. 1a and Supplementary Fig. 1), we first compared photocrosslinking methods^{28–30} using stable Flag-haemagglutinin (HA)-tagged FMRP isoform 7 in HEK293 cells (Fig. 1b), as these cells and human brain share 90% of

expressed genes according to a comparison of existing RNA-seq data sets^{31–33} (Supplementary Fig. 2). The difference between *FMR1* levels in the experimental system and the brain is 1.3-fold, as calculated using RNA-seq data and the quantified expression of *FMR1* in our stable cells. We found that 4-thiouridine (4SU) photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) provided the highest yield of crosslinked RNAs, and this approach was used for all FMR1-family proteins (Fig. 1c). Complementary DNA libraries were generated after PAR-CLIP and Illumina-sequenced (Supplementary Table 1). Genome-aligned reads were grouped by PARalyzer³⁴ to identify segments of RNA that represented peaks of T-to-C conversion, termed binding sites. PARalyzer separated closely spaced binding sites connected by overlapping reads and yielded a median RNA segment length of 33 nucleotides (Supplementary Fig. 3). FMRP isoforms 1 and 7 bound to approximately 80,000–100,000 sites, of which >85% mapped to ~6,000 mRNAs (Supplementary Tables 1, 2 and <http://fmrp.rockefeller.edu>). FXR1P and FXR2P protein binding sites are comprised within FMRP binding sites, with an overlap of 95% (Supplementary Table 3).

Nearly all mRNA-binding sites were located in exons (>90%) (Fig. 2a) and distributed between coding sequence (CDS) and 3' untranslated regions (UTRs) (>95%, total), with slightly more CDS sites, similar to distributions seen for other cytoplasmic RNA-binding proteins²⁸. The computational sequence analysis method cERMIT³⁵ revealed two key RREs, ACUK and WGGA (in which K = G or U and W = A or U) (Fig. 2b and Supplementary Fig. 4). Together, ACUK and WGGA RREs were found in ≥50% of mRNA-binding sites in isoforms 1 and 7, occurring exclusively or together within the same binding site (Fig. 2c). Remaining binding sites typically contained close derivatives of either RRE.

¹Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, New York 10065, USA. ²Institute for Genome Sciences and Policy, Duke University, Durham, North Carolina 27708, USA. ³Program for Early and Recurrent Pregnancy Loss, Department of Obstetrics & Gynecology and Women's Health, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ⁴Genomics Resource Center, The Rockefeller University, New York, New York 10065, USA. [†]Present address: The Berlin Institute for Medical Systems Biology, Max Delbrück Center, 13125 Berlin-Buch, Germany (N.M. and U.O.).

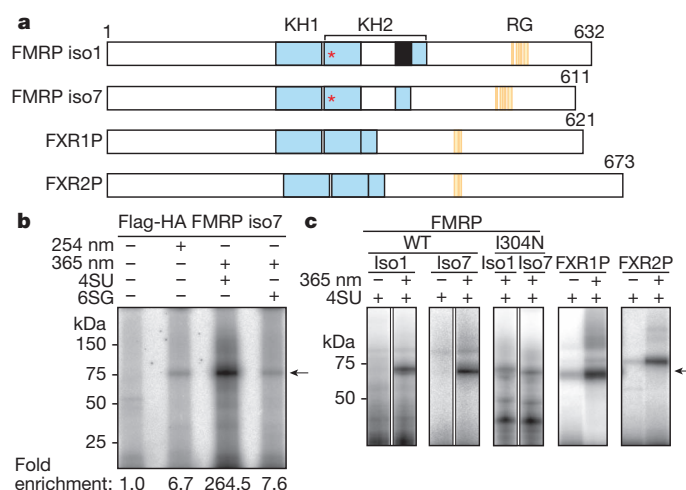


Figure 1 | PAR-CLIP of FMR1-family proteins. **a**, FMR1-family proteins comprise two type-I KH domains (cyan). FMRP isoforms (iso) 1 and 7 vary by the presence of exon 12 (black) within KH2. The I304N mutation (red asterisk) is located within the KH2 domain. The arginine-glycine-rich region (RG; orange bars) is also implicated in RNA binding. The lengths of proteins in amino acids are indicated. We established stable inducible cell lines expressing Flag-HA epitope-tagged wild-type and I304N mutants of FMRP (isoforms 1 and 7), and its paralogues FXR1P and FXR2P (ref. 47). **b**, RNA-FMRP crosslinking comparing CLIP (254 nm) to 4SU or 6-thioguanosine (6SG) PAR-CLIP (365 nm). RNA-radiolabelled Flag immunoprecipitates of lysates from crosslinked HEK293 cells expressing Flag-HA-tagged FMRP isoform 7 were separated by SDS-polyacrylamide gel electrophoresis (PAGE). The migrations of protein mass standards are indicated. Enrichment of radiolabelled RNA covalently bound to Flag-HA-FMRP (arrow) was determined after normalizing by western blot analysis (not shown). **c**, 4SU PAR-CLIP of FMR1-family proteins analogous to **b**.

Characterization of the RREs of FMRP

We performed electrophoretic mobility shift assays (EMSAs) to test RNAs representing FMRP binding sites using recombinant human FMRP purified from *Spodoptera frugiperda* 9 (Sf9) cells (Supplementary Figs 5 and 6). FMRP target sites were selected on the basis of whether they contained ACUK, WGGA or mixed (ACUK and WGGA) RREs (Supplementary Table 4). Testing RNAs of various lengths, we found that oligonucleotide lengths of ≥ 45 nucleotides were required to observe gel shifts and reach dissociation constants below $0.5 \mu\text{M}$, suggesting that RNA backbone contacts outside of the RREs contribute to the association *in vitro*. WGGA-containing RNAs exhibited the widest range and strongest affinities, generally correlating with the number of RREs within a PARalyzer-defined binding site. An RNA segment containing nine WGGA bound almost two orders of magnitude tighter than those containing one WGGA, whereas binding of ACUK-containing RNAs varied only fivefold. EMSAs using RNAs representing target sites within *PPP2CA* and *UBE3A* are shown in Fig. 3a.

We also tested sites in *APP* and *FMR1* mRNA (Supplementary Fig. 7), two previously identified mRNA targets. *APP* was originally discovered as an FMRP target on the basis of a predicted G-quartet region⁷ although the actual site was subsequently identified *in vitro* at an upstream segment³⁶, which was identified here (*APP* site 1) as containing WGGA. FMRP targets its own mRNA, although this association was only observed *in vitro*^{5,6,37} and in immunoprecipitates³⁸.

Recombinant I304N isoform 1 FMRP showed an \sim tenfold average decrease in its affinity towards ACUK-containing RNAs compared to wild-type FMRP isoform 1 (Fig. 3b and Supplementary Table 5). We characterized binding to wild-type and mutant RNA sequences derived from an *NFI* mixed RRE site (Fig. 3c). I304N FMRP bound wild-type and *NFI* ACUK(−) RNAs similarly, whereas wild-type FMRP showed a twofold reduction in affinity for *NFI* ACUK(−)

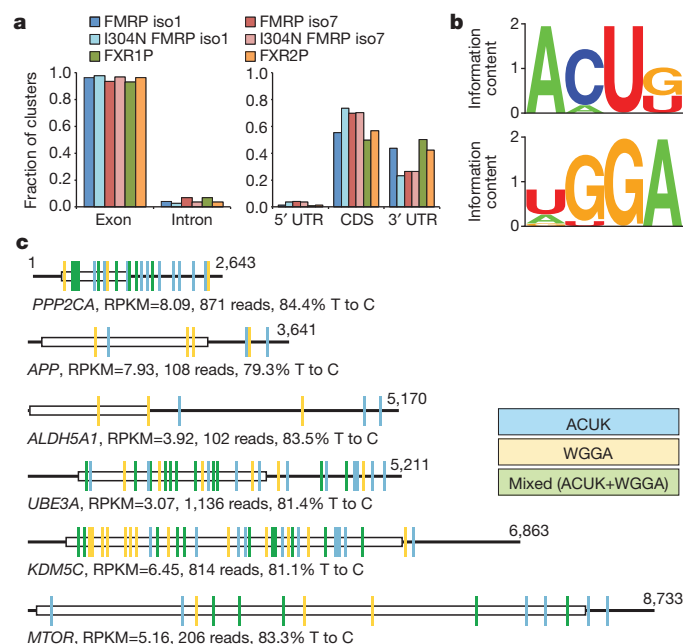


Figure 2 | Analysis of FMR1-family protein mRNA-binding sites.

a, Distribution of binding sites within mRNA targets of the FMR1 protein family. **b**, Two key RREs were inferred from FMRP isoform 1 and 7 binding sites: top, ACUK; bottom, WGGA. **c**, Distribution of FMRP binding sites, colour-coded on the basis of cERMIT-inferred RREs, across representative targets. Open boxes and thick lines indicate CDS and UTRs, respectively. Numbers indicate nucleotide number. RPKM, reads per kilobase of mature transcript per million mapped reads.

RNA. In addition, wild-type FMRP binding curves for *NFI* ACUK, WGGA(−) RNA were similar to I304N FMRP binding curves for *NFI* WGGA(−) RNA. Our results indicate that the KH2 domain associates with the ACUK RRE. As FMRP associates with mRNAs at more than one binding site, its association at multiple sites will affect the final regulatory effect. We compared the distribution of RREs in I304N FMRP versus wild-type FMRP PAR-CLIPs and found a transcriptome-wide depletion in the recovery of sequence reads for ACUK binding sites for both I304N isoforms, consistent with the biochemistry (Fig. 3d). Interestingly, the RRE fractional distribution of I304N FMRP isoform 1 was similar to wild-type FMRP isoform 7, suggesting that alteration of the KH2 domain by mutation or exon insertion affects binding-site selectivity. Taken together, the biochemical and PAR-CLIP results with I304N FMRP indicate that we identified the natural target sites of the FMRP KH2 domain.

Ranked enrichment of the mRNA targets of FMRP

To rank FMRP targets, we measured their enrichment by ribonucleoprotein immunoprecipitation followed by microarray analysis (RIP-chip), which would indicate stable interactions. In total, 3,593 PAR-CLIP-identified targets showed enrichment by RIP-chip, of which 939 genes were two- to sixfold enriched; 646 transcripts were twofold enriched but not identified as PAR-CLIP targets. We used binding-site information obtained by PAR-CLIP to infer the salient features for stable association in RIP-chip (Fig. 4, Supplementary Fig. 8 and Supplementary Table 6). Increasing frequency of WGGA- and ACUK-containing elements led to greater RIP-chip enrichment, in agreement with *in vitro* affinity measurements. On average, top targets contained more RRE binding sites (18 per transcript) compared to the least-enriched targets (13 per transcript). Top-ranked targets had a CDS/3' UTR binding-site ratio of 3.7/1 compared to bottom-ranked targets, which had a 1/2 ratio. Transcripts with ACUK/WGGA RRE ratios < 1 were the most statistically significantly ($P = 8.73 \times 10^{-9}$) enriched population. Importantly, we

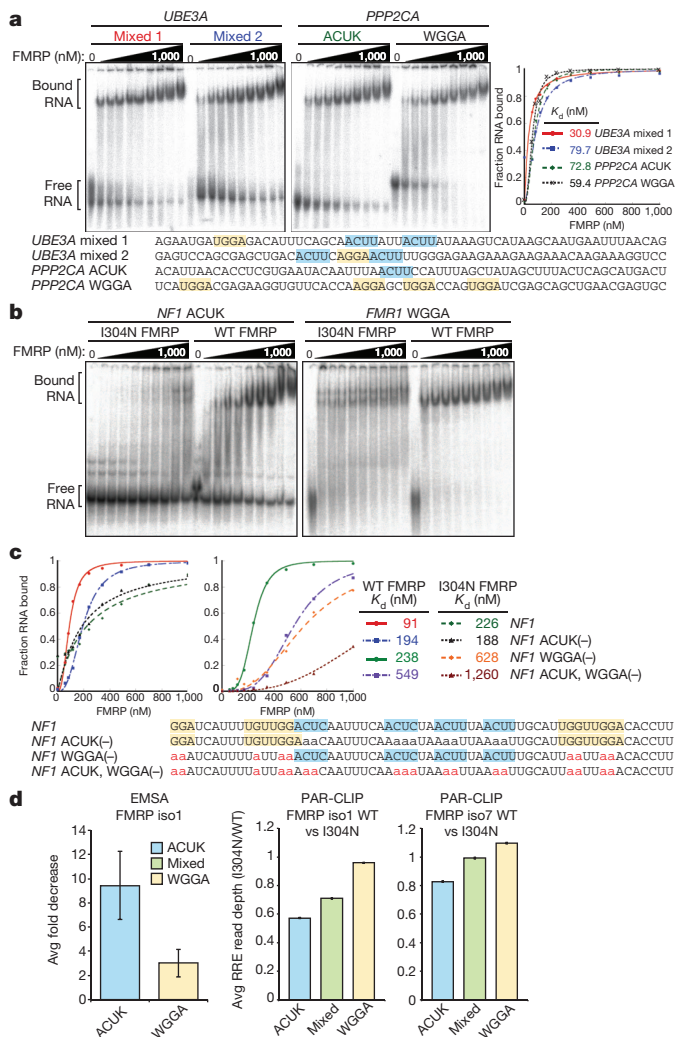


Figure 3 | RNA-binding assays using natural FMRP target sites containing ACUK and WGGA RREs, and the effect of a KH2 mutation to its target RNA spectrum. **a**, Top left, EMSAs of RNAs representing *UBE3A* or *PPP2CA* binding sites containing various RREs. Top right, binding curves and dissociation constants (K_d) are shown. Bottom, the sequences of the RNAs are provided, with WGGA (yellow) and ACUK RREs (cyan) highlighted. **b**, Effect of the KH2 mutation in FMRP on target sites containing ACUK versus WGGA RREs. The RNA affinities of wild-type and I304N FMRP isoform 1 were compared using binding sites in *NF1* (ACUK) and *FMRI* (WGGA). **c**, Binding curves of wild-type and I304N FMRP for an RNA segment representing a mixed RRE binding site in *NF1*, and several mutant sequence versions (ACUK(-), WGGA(-) and ACUK, WGGA(-)). **d**, Comparison of FMRP isoform 1 affinity for RRE type in EMSAs and FMRP isoforms 1 and 7 wild-type and I304N PAR-CLIP libraries. Error bars in EMSA summary (left) represent s.d., $n = 9$ (ACUK) or 8 (WGGA). The ratio of sequence reads aligned to each RRE binding site was calculated between wild-type and I304N FMRP PAR-CLIP libraries. The average sequence-depth ratio of wild-type over I304N binding site, per RRE type, are shown. Error bars in the read-depth analyses (middle and right) represent the average minimum and maximum values across all sub-sampled mutant libraries ($n = 14$ and 26 for isoforms 1 and 7, respectively).

identified 93 genes independently implicated in ASD among the highly enriched FMRP targets (Fig. 4d and Supplementary Table 7), which is notable given the relationship between FXS and ASD^{20,21,39}.

Enrichment of RNAs in RIP-chip depends on the saturation of target sites with FMRP. To estimate the yield of saturation, protein copy number, target mRNA copies and the number of binding sites within those transcripts have to be accounted. We determined endogenous and co-expressed Flag-HA-tagged FMR1 protein copy numbers to be approximately 60,000 and 70,000 molecules per cell,

respectively, using quantitative western blotting and reference recombinant protein. Each cell contains 20 pg total RNA, 4% of which is mRNA and representing approximately 1 million molecules. Considering their relative abundance on the basis of HEK293 RNA-seq, we estimate that an equal distribution of FMRP would occupy 6% of binding sites per cell among its 6,000 target transcripts, or up to 20% if FXR1P and FXR2P protein estimates are included. Because PAR-CLIP-identified targets had varying enrichment, their association with FMRP is not uniform. The ~3,500 transcripts enriched in RIP-chip are estimated to have at least 18% FMRP binding-site occupancy, whereas the 900 top-ranked genes (twofold enriched) potentially exhibit 78%. The presence of multiple binding sites within targets suggests that multiple FMR1-family proteins bind each transcript to influence their regulation.

Regulatory impact of FMRP on its mRNA targets

To assess the effect of FMRP binding sites on mRNA stability, small interfering (siRNA) knockdown of *FMRI* or the FMR1 family was performed and mRNA-expression profiles were analysed by microarray. We found no evidence for FMRP affecting target mRNA abundance (data not shown).

A panel of FMRP targets was selected on the basis of enrichment in RIP-chip, low-to-intermediate expression in RNA-seq, similar abundance in the human brain (using published microarray³¹ and RNA-seq data sets^{32,33}) and with documented neurological and human disease relevance, and then analysed by quantitative western blot to determine protein levels as a function of FMRP expression (Fig. 4e). FXR2P, HUWE1, KDM5C and MTOR protein levels, among others, showed up to 30% reduction in protein levels upon expression of FMRP in HEK293. We analysed lysates prepared from human post-mortem brains. Four FXS brains (Supplementary Fig. 9) were available with age/sex/anatomic-matched controls from prefrontal cortical, hippocampal and cerebellar regions. Although only four out of eight antibodies yielded quantifiable bands in brain lysates, we observed a general trend of elevated target protein expression levels in FXS brains. This is the inverse of FMRP-overexpression effects in HEK293, and consistent with FMRP affecting the protein levels of its mRNA targets.

The mRNA targets identified here are from a human transcriptome in which most genes are comparably expressed in the human brain (Supplementary Fig. 2). We discovered ASD-related and numerous other genes implicated in neuronal disorders associated with FXS and validated representatives by EMSA, RIP-chip and immunoblot. We found genes involved in Angelman, Prader-Willi, Rett, and Cornelia de Lange syndromes. Interestingly, the ASD and Angelman syndrome-associated gene *UBE3A* ubiquitinates ARC and SACS⁴⁰; ARC is a well-known target and here we identify SACS as a targeted transcript. These findings potentially provide the molecular link to tie together elements of clinically overlapping disorders, principally setting a molecular target framework for characterizing the connections between FXS and its associated phenotypes.

Although fragile-X-related diseases are primarily considered to be CNS disorders, at least two other target organs are affected: the testes and ovaries. We reasoned that changes in *FMRI* expression lead to dysregulation of largely overlapping sets of targets shared across all affected organs. Thus, dysregulated genes and pathways in the brain might also contribute to phenotypes in the testes and ovaries. We therefore examined the ovaries of *Fmr1*^{-/-} mice⁴¹ as CNS and macroorchidism phenotypes had been reported, yet ovary development had largely been under-investigated. Ovaries from *Fmr1*^{-/-} mice were markedly larger by 3 weeks post-birth compared to wild-type controls (Fig. 5a, b). At 12 and 18 weeks post-birth, knockout ovaries were 22% and 72% larger by mass compared to age-matched controls, respectively. Notably, we found increased protein levels of Tsc2, Sash1 and Mtor (Fig. 5c). As a role for the Mtor pathway in the regulation of ovarian development has been previously identified, it is tempting to

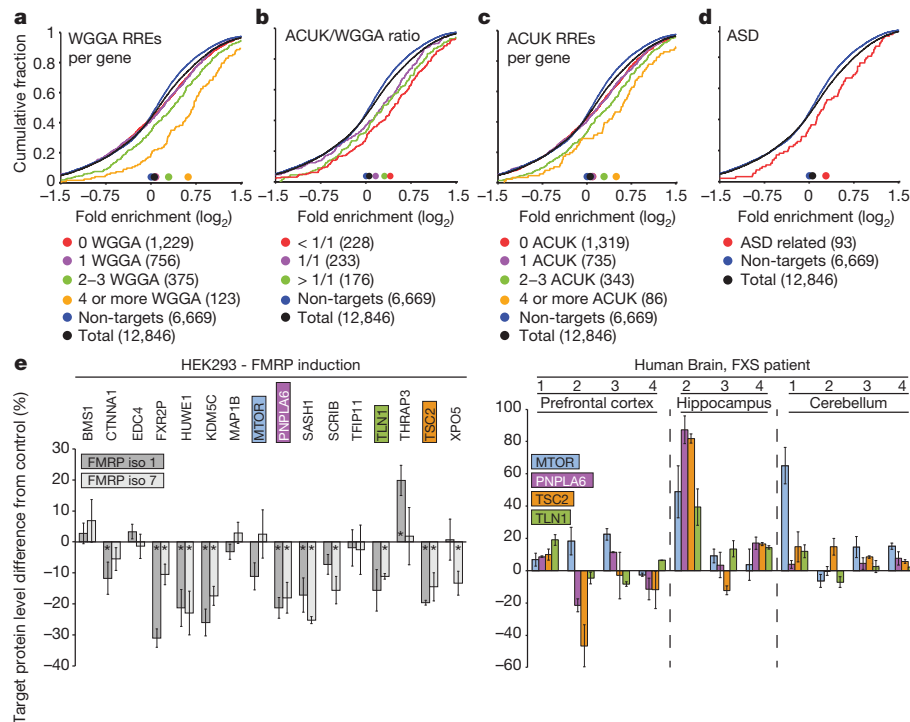


Figure 4 | RRE-dependent enrichment criteria for FMRP association with mRNAs. **a–d**, Cumulative distribution fraction plots of FMRP targets on the basis of indicated criteria. Transcripts were grouped and colour-coded on the basis of indicated bins. Non-targets are mRNA transcripts with zero PAR-CLIP binding sites, although detectable in the array; total denotes the sum of non-targets and PAR-CLIP-identified targets detectable by RIP-chip. RIP-chip experiments were performed using Flag-HA-FMRP isoform 1 (**a**). Enrichment of 93 PAR-CLIP identified ASD-related target genes (**d**). **e**, Immunoblot densitometry analysis of top-ranking FMRP targets from RIP-chip and

PAR-CLIP analyses in HEK293 cells and human brain. In cell culture, target-protein expression differences of indicated proteins were determined upon induction of FMRP isoform 1 or 7 expression. Similarly, relative protein expression was measured using lysates prepared from indicated brain regions of four FXS patients, compared to age/sex/anatomic-matched controls. Error bars denote s.e.m.; $n = 2–11$ (depending on protein measured and whether the sample was HEK293 or brain lysate). PABPC1 protein level served as a loading and ratio control as it was a gene with PAR-CLIP binding sites but showed no RIP-chip enrichment ($-0.10 \log_2$ -fold enrichment). $*P < 0.05$.

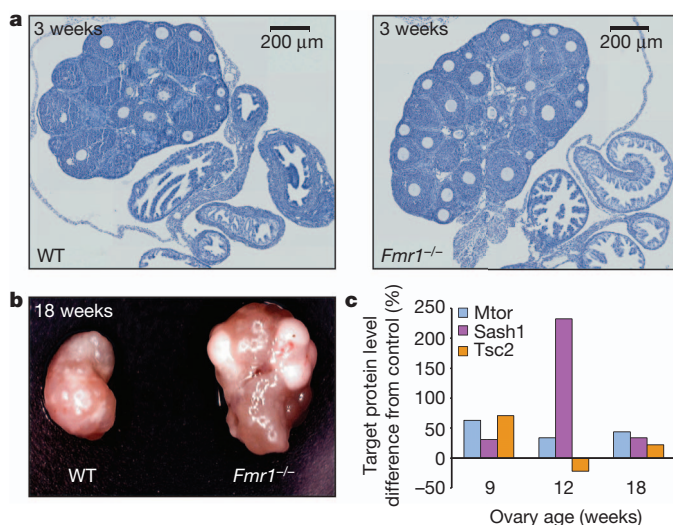


Figure 5 | Ovarian phenotype in *Fmr1*^{-/-} mice. **a–c**, Ovaries from wild-type and *Fmr1*^{-/-} female mice were collected at 3, 9, 12 and 18 weeks and processed for histological (**a**), morphological (**b**) and quantitative western analyses (**c**). By 3 weeks of age, histological staining (haematoxylin) of sectioned ovaries showed a greater than expected number of follicles compared to wild type (**a**). Ovaries from 18-week-old *Fmr1*^{-/-} mice are larger than those from wild type and exhibit prominent cysts consistent with corpus luteal development (**b**). Lysates were prepared from 9-, 12- and 18-week-old ovaries from two different wild-type and knockout mice each, and analysed by quantitative western blot using Mtor, Sash1 and Tsc2 antibodies. As for human samples, Pabpc1 was used for normalization (**c**).

conclude that increased activity, in the absence of *Fmr1* expression, contributes to enlarged ovaries histologically consistent with precocious follicular development. These observations suggest that *Fmr1* knockout mice have the potential to model FXS-related premature ovarian insufficiency, in which it remains unclear whether elevated *FMR1* mRNA or the observed reduction of FMRP protein itself is causative in female carriers affected by this disease.

Increased signal-transduction activity of the Mtor pathway^{42,43} was reported in *Fmr1* knockout mice, and was attributed to *Fmrp* targeting of *Pik3ca* mRNA. Indeed, PAR-CLIP identified several FMRP binding sites within the human *PIK3CA* transcript. However, we find that it is a less-enriched RIP-chip target compared to *MTOR* and *TSC2*, whose protein levels appear regulated in an FMRP-dependent manner. Interestingly, recent evidence demonstrated that *Tsc* mutant mice^{44–46}, which have increased Mtor activity, had impaired mGluR-induced long-term synaptic depression and protein synthesis compared to *Fmr1*^{-/-} mice; crossing *Tsc2*^{+/-} with *Fmr1*^{-/-} mice corrected the phenotypes⁴⁴. Given our results it is likely that the decreased protein synthesis observed in *Tsc2*^{+/-} mice is due to the Mtor pathway inducing increased FMRP regulatory activity, a consequence which can be alleviated through the removal of FMRP itself in *Tsc2*^{+/-} *Fmr1*^{-/-} mice. However, it should be noted that FMRP can also associate with transcripts of ERK pathway components. Therapeutic targeting of the MTOR pathway has become an important goal, but must be further guided by additional functional analysis, particularly of FMRP targets upstream and downstream of MTOR and interconnected signalling pathways (Supplementary Fig. 10). Combined, our validation work in *Fmr1* knockout mouse ovaries and in human brain demonstrate that the effect of FMRP binding to specific target genes identified in cell culture is extensible to physiologically relevant contexts.

METHODS SUMMARY

Methods are described in greater detail in Supplementary Information. FVB129P2 fragile X mice were a gift from S. Zukin. Gateway plasmids (Invitrogen) generated in this study will be deposited in <http://www.addgene.org>. FLPIn TREx HEK293 (Invitrogen) inducible stable cell lines were generated per manufacturer's instructions. The titres, source and use of antibodies used in this study are listed in Supplementary Information. PAR-CLIP was performed essentially as described, except that the second RNase T1 digestion was omitted following the immunoprecipitation. Recombinant wild-type and mutant FMRP isoform 1 proteins were purified using a baculovirus expression system (Invitrogen). EMSAs and western blots were quantified using ImageGauge (Fuji). Parameters of computation analyses are described in Supplementary Information and in the relevant sections at <http://fmrp.rockefeller.edu/>. Relevant data sets, including raw data, are available at <http://fmrp.rockefeller.edu/> and gene expression omnibus (GSE39686).

Full Methods and any associated references are available in the online version of the paper.

Received 29 June; accepted 2 November 2012.

Published online 12 December 2012.

- Verkerk, A. J. *et al.* Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65**, 905–914 (1991).
- Kremer, E. J. *et al.* Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)_n. *Science* **252**, 1711–1714 (1991).
- Vincent, A. *et al.* Abnormal pattern detected in fragile-X patients by pulsed-field gel electrophoresis. *Nature* **349**, 624–626 (1991).
- Siomi, H., Siomi, M. C., Nussbaum, R. L. & Dreyfuss, G. The protein product of the fragile X gene, *FMR1*, has characteristics of an RNA-binding protein. *Cell* **74**, 291–298 (1993).
- Ashley, C. T. J., Wilkinson, K. D., Reines, D. & Warren, S. T. FMR1 protein: conserved RNP family domains and selective RNA binding. *Science* **262**, 563–566 (1993).
- Brown, V. *et al.* Purified recombinant Fmrp exhibits selective RNA binding as an intrinsic property of the fragile X mental retardation protein. *J. Biol. Chem.* **273**, 15521–15527 (1998).
- Darnell, J. C. *et al.* Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell* **107**, 489–499 (2001).
- Ashley, C. T. *et al.* Human and murine *FMR-1*: alternative splicing and translational initiation downstream of the CGG-repeat. *Nature Genet.* **4**, 244–251 (1993).
- Verkerk, A. J. *et al.* Alternative splicing in the fragile X gene *FMR1*. *Hum. Mol. Genet.* **2**, 399–404 (1993).
- De Boule, K. *et al.* A point mutation in the *FMR-1* gene associated with fragile X mental retardation. *Nature Genet.* **3**, 31–35 (1993).
- Siomi, H., Choi, M., Siomi, M. C., Nussbaum, R. L. & Dreyfuss, G. Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. *Cell* **77**, 33–39 (1994).
- Feng, Y. *et al.* FMRP associates with polyribosomes as an mRNP, and the 1304N mutation of severe fragile X syndrome abolishes this association. *Mol. Cell* **1**, 109–118 (1997).
- Fridell, R. A., Benson, R. E., Hua, J., Bogerd, H. P. & Cullen, B. R. A nuclear role for the Fragile X mental retardation protein. *EMBO J.* **15**, 5408–5414 (1996).
- Ling, S. C., Fahrner, P. S., Greenough, W. T. & Gelfand, V. I. Transport of Drosophila fragile X mental retardation protein-containing ribonucleoprotein granules by kinesin-1 and cytoplasmic dynein. *Proc. Natl Acad. Sci. USA* **101**, 17428–17433 (2004).
- Davidovic, L. *et al.* The fragile X mental retardation protein is a molecular adaptor between the neurospecific KIF3C kinesin and dendritic RNA granules. *Hum. Mol. Genet.* **16**, 3047–3058 (2007).
- Dicthenberg, J. B., Swanger, S. A., Antar, L. N., Singer, R. H. & Bassell, G. J. A direct role for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. *Dev. Cell* **14**, 926–939 (2008).
- Corbin, F. *et al.* The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes. *Hum. Mol. Genet.* **6**, 1465–1472 (1997).
- Oostra, B. A. & Willemsen, R. *FMR1*: a gene with three faces. *Biochim. Biophys. Acta* **1790**, 467–477 (2009).
- De Rubéis, S., Fernández, E., Buzzi, A., Di Marino, D. & Bagni, C. Molecular and cellular aspects of mental retardation in the Fragile X syndrome: from gene mutation/s to spine dysmorphogenesis. *Adv. Exp. Med. Biol.* **970**, 517–551 (2012).
- Gross, C., Berry-Kravis, E. M. & Bassell, G. J. Therapeutic strategies in fragile X syndrome: dysregulated mGluR signaling and beyond. *Neuropsychopharmacology* **37**, 178–195 (2012).
- Hagerman, R., Lauterborn, J., Au, J. & Berry-Kravis, E. Fragile X syndrome and targeted treatment trials. *Results Probl. Cell Differ.* **54**, 297–335 (2012).
- Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
- Krueger, D. D. & Bear, M. F. Toward fulfilling the promise of molecular medicine in fragile X syndrome. *Annu. Rev. Med.* **62**, 411–429 (2011).
- Pfeiffer, B. E. & Huber, K. M. The state of synapses in fragile X syndrome. *Neuroscientist* **15**, 549–567 (2009).
- Belmonte, M. K. & Bourgeron, T. Fragile X syndrome and autism at the intersection of genetic and neural networks. *Nature Neurosci.* **9**, 1221–1225 (2006).
- Brown, V. *et al.* Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* **107**, 477–487 (2001).
- Bassell, G. J. & Warren, S. T. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* **60**, 201–214 (2008).
- Hafner, M. *et al.* Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).
- Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* **456**, 464–469 (2008).
- Ascano, M., Hafner, M., Cekan, P., Gerstberger, S. & Tuschl, T. Identification of RNA-protein interaction networks using PAR-CLIP. *Wiley Interdiscip. Rev. RNA* **3**, 159–177 (2012).
- Su, A. I. *et al.* A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl Acad. Sci. USA* **101**, 6062–6067 (2004).
- Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**, 470–476 (2008).
- Kishore, S. *et al.* A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nature Methods* **8**, 559–564 (2011).
- Corcoran, D. L. *et al.* PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. *Genome Biol.* **12**, R79 (2011).
- Georgiev, S. *et al.* Evidence-ranked motif identification. *Genome Biol.* **11**, R19 (2010).
- Westmark, C. J. & Malter, J. S. FMRP mediates mGluR5-dependent translation of amyloid precursor protein. *PLoS Biol.* **5**, e52 (2007).
- Schaeffer, C. *et al.* The fragile X mental retardation protein binds specifically to its mRNA via a purine quartet motif. *EMBO J.* **20**, 4803–4813 (2001).
- Ceman, S., Brown, V. & Warren, S. T. Isolation of an FMRP-associated messenger ribonucleoprotein particle and identification of nucleolin and the fragile X-related proteins as components of the complex. *Mol. Cell. Biol.* **19**, 7925–7932 (1999).
- Bhakar, A. L., Dölen, G. & Bear, M. F. The pathophysiology of fragile X (and what it teaches us about synapses). *Annu. Rev. Neurosci.* **35**, 417–443 (2012).
- Greer, P. L. *et al.* The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. *Cell* **140**, 704–716 (2010).
- Bakker, C. E. *et al.* Fmr1 knockout mice: a model to study fragile X mental retardation. The Dutch-Belgian Fragile X Consortium. *Cell* **78**, 23–33 (1994).
- Gross, C. *et al.* Excess phosphoinositide 3-kinase subunit synthesis and activity as a novel therapeutic target in fragile X syndrome. *J. Neurosci.* **30**, 10624–10638 (2010).
- Sharma, A. *et al.* Dysregulation of mTOR signaling in fragile X syndrome. *J. Neurosci.* **30**, 694–702 (2010).
- Auerbach, B. D., Osterweil, E. K. & Bear, M. F. Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* **480**, 63–68 (2011).
- Bateup, H. S., Takasaki, K. T., Saulnier, J. L., Deneff, C. L. & Sabatini, B. L. Loss of Tsc1 *in vivo* impairs hippocampal mGluR-LTD and increases excitatory synaptic function. *J. Neurosci.* **31**, 8862–8869 (2011).
- Chévere-Torres, I. *et al.* Metabotropic glutamate receptor-dependent long-term depression is impaired due to elevated ERK signaling in the ΔRG mouse model of tuberous sclerosis complex. *Neurobiol. Dis.* **45**, 1101–1110 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements Human tissue was obtained from the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland, Baltimore. We would like to thank the following members of the Tuschl laboratory for their support and assistance: G. Wardle, N. Renwick and I. Ben-Dov. We would like to thank J. Keene for his advice throughout the project. We would like to acknowledge M. Khorshid, L. Burger and M. Zavolan for analysing PAR-CLIP data at the initial stages of the project and discussions. We would like to thank the Memorial Sloan-Kettering Cancer Center in-situ core for their assistance with the mouse histology. Finally, we would like to express our appreciation to all members of the Tuschl laboratory for their assistance and collegiality. This work was supported, in part, by the following agencies: NIH/NCRR/RU CCTS (M.A., UL1RR024143), NSF (U.O., MCB-0822033), Simons Foundation Autism Research Initiative (T.T., CEN5300891) and NIH (T.T., R01 MH080442; Z.W., K08 HD068546).

Author Contributions M.A. designed, executed, supervised and interpreted experiments. N.M., P.B. and D.L.C. carried out the sequence alignment, annotation and PARalyzer pipeline. N.M. and P.B. performed the computational analysis on the RIP-chip. M.A., J.B.M. and J.D.N. purified FMRP proteins, performed the EMSAs and carried out the quantitative western blots and analyses. M.A. and M.M. performed the RIP-chips. S.D. assisted in the Illumina sequencing of all PAR-CLIP libraries. M.H. helped in the initial PAR-CLIP experiments. C.L. and Z.W. carried out and analysed mouse experiments. U.O. supervised computational efforts. T.T. supervised and helped in the design of experiments. M.A. and T.T. wrote the manuscript.

Author Information Data sets have been submitted to gene expression omnibus (GEO) under the accession code GSE39686. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.T. (ttuschl@rockefeller.edu) or U.O. (uwe.ohler@duke.edu).

METHODS

Animals. Wild-type and knockout FVB129P2 fragile X mice were a gift from S. Zukin and were maintained under standard conditions. Female mice, 21–30 days old, were killed by cervical dislocation and the ovaries were aseptically removed. Dissection of the ovary from the fat pad and the bursa was done using a stereomicroscope at $\times 2$ magnification. Excised ovaries were gently washed using PBS, weighed on an analytical balance and used either for paraffin embedding and immunohistological studies or for protein extraction. All procedures were approved by the Rockefeller University Institutional Animal Care and Use Committee (IACUC).

Histology. Mouse ovaries were fixed in 4% paraformaldehyde overnight at 4 °C, embedded in paraffin, cut into 4- μ m sections, and applied to Superfrost/Plus slides (Fisher Scientific). Haematoxylin staining was performed by the Molecular Cytogenetic core facility of Memorial Sloan-Kettering Cancer Center. **Plasmids.** Plasmids pENTR4-FMRP isoform 7, -FXR1P and -FXR2P were generated by restriction enzyme digestion and ligation of the respective PCR products into pENTR4 (Invitrogen). pENTR4-FMRP isoform 1 was obtained by restriction enzyme digestion and ligation of the 525-base pair PCR product, corresponding to exon 12 and a portion of exons 11 and 13, into the pENTR4-FMRP isoform 7 plasmid. pENTR4-I304N FMRP mutant isoforms were generated by QuikChange (Agilent) according to the manufacturer's instructions. The pENTR4 vectors were subsequently recombined into the pFRT/TO/Flag/HA-DEST destination vectors using Gateway LR recombinase according to manufacturer's protocol (Invitrogen). All plasmids generated in this study will be deposited at <http://www.Addgene.org>.

Cell lines and culture. Generation and maintenance of stable FMR1-family proteins expressed in FlpIn TREx HEK293 cell lines were performed as described in ref. 47. Expression of FMR1-family proteins was induced by supplementing medium with 1 μ g ml⁻¹ doxycycline for 16 h, prior to any analysis.

Oligodeoxynucleotides and DNA subcloning. The following oligodeoxynucleotide primers were used for PCR and subcloning (restriction sites are underlined): FMR1 isoform 7 (accession version NM_002024.4): 5'-TACCATTGGAGGAGCTGGTGGTGAAG-3', 5'-CGGAATTCGTTTAGGGTACTCCATTCACGA-3'; FXR1 (accession version NM_005087.2): 5'-TACCATTGGCGGACGTGACGGTGA-3', 5'-CGGAATTCGTTTATGAAACACCATTCAGGACTGC-3'; FXR2 (accession version NM_004860.2): 5'-TAGTCGACATGGGCGGCCTGGCCTC-3', 5'-CGGAATTCGTTTATGAAACCCATTCACCATAC-3'. To obtain the FMR1 isoform 1-specific DNA segment, the following oligodeoxynucleotide primers were used to PCR-amplify FMR1 exon 12 from cDNA generated from HEK293 cells (SMART RACE, Clontech): 5'-GGAGAGGATCAGGATGCAGTG-3', 5'-CTGCTCATCAATTTGTAATCTC-3'. I304N FMRP mutant isoforms were generated using the following primers: 5'-GTAATAGGAAAAATGGAAAGCTGAATCAGGAGATTTGGGACAAGTC-3' and 5'-GACTTGTCACAACTCCTGATTCAGCTTTCCATTTTCTATTAC-3', in which the mutation is indicated in lowercase letters.

Antibodies. The following mouse monoclonal antibodies (titre, vendor, catalogue no.) were used for immunoblots: anti-HA.11 (1:1,000, Covance, MMS-101P), anti-FMR1 (1:1,000, Millipore, MAB2160), anti-BMS1 (1:100, Abcam, AB50698), anti-CTNNA1 (1:1,000, Sigma-Aldrich, SAB1402163), anti-FXR2P (1:2,000, Millipore, 05-1214), anti-MAP1B (1:500, Abcam, AB11266), anti-TLN (1:1,000, Abcam, AB57758), anti-XPO5 (1:200, Signal-Aldrich, WH0057510M1). The following rabbit antibodies were used for immunoblots: anti-EDC4 (1:2,000, Abcam, AB72408), anti-HUWE1 (1:1,000, Bethyl, A300-486A), anti-KDM5C (1:250, Abcam, AB34718), anti-MTOR (1:1,000, Cell Signaling, 2972), anti-PNPLA6 (1:1,000, Protein Tech, 14261-1), anti-SASH1 (1:2,000, Novus, NBP1-26650), anti-SCRIB (1:1,000, Cell Signaling, 4475), anti-TFIP11 (1:2,000, Bethyl, A302-548A), anti-THRAP3 (1:2,000, Bethyl, A300-956A), anti-TSC2 (1:1,000, Cell Signaling, 3612), anti-PABPC1 (1:1,000, Cell Signaling, 3505). The following mouse monoclonal antibodies were used for EMSA supershift analyses: anti-Flag M2 (Sigma-Aldrich, F1804), anti-penta-His (Qiagen, 34,660), anti- β -tubulin (Sigma-Aldrich, T4026), anti-FMR1 (Millipore, MAB2160). Horseradish-peroxidase-conjugated polyclonal goat secondary antibodies raised against rabbit or mouse immunoglobulin (Dako, P0448 and P0447) were used at a titre of 1:3,000, in conjunction with the appropriate species primary antibody, for immunoblot analyses.

Immunoblotting and quantitative chemiluminescence analyses. Protein samples were prepared by addition of $\times 4$ sample buffer (200 mM Tris-HCl, pH 6.8, 200 mM dithiothreitol (DTT), 8% SDS, 24% glycerol, 0.04% bromophenol blue) and incubation for 5 min at 95 °C. Samples were separated by SDS-PAGE (6–12%, Bio-Rad Mini Protean II, 7.3 \times 10.2 cm) for approximately 1 h at 25 mA per gel, using standard SDS Tris base-glycine running buffer. Alternatively, protein samples were loaded onto pre-cast 4–12% linear-gradient SDS Tris-glycine or Tris-acetate gels (Invitrogen) and run according to manufacturer's

instructions. After electrophoresis, proteins were blotted onto nitrocellulose membranes (Hybond-ECL, GE Life Science) pre-wetted in transfer buffer (25 mM Tris base, 190 mM Glycine, 20% MeOH, 0.05% SDS) and semi-dry transferred (Bio-Rad) at 300 mA for 1.5 h. Protein membranes were taken through a standard immunoblot protocol followed by enhanced chemiluminescent detection (Crescendo ECL, Millipore) using a Lumimager (Fuji, LAS-3000). Quantification of bands was performed using ImageGauge v4.1 (Fuji) and Excel (Microsoft) software. The band intensities of each protein of interest (Figs 4 and 5) for each set of replicate samples (control, FMRP isoform 1, etc.) were measured, and the means and standard errors were calculated. The band intensities for reference PABPC1 protein expression were similarly quantified and then used to correct for loading differences. Migration of PABPC1 by SDS-PAGE was sufficiently different compared to the test proteins; therefore immunoblots against test proteins and PABPC1 were done in parallel from the lanes of the same nitrocellulose membranes by mixing primary antibodies during incubation. Prior to this procedure, each primary antibody was tested singularly. Fold differences were calculated by dividing the mean intensity of the protein of interest in experimental samples over the mean intensity of that protein of interest in Control samples after loading normalization (with PABPC1).

PAR-CLIP cDNA library generation. Concurrent with the addition of doxycycline, 100 μ M 4SU or 6SG was added to the cells and incubated for 16 h, prior to 365 nm ultraviolet irradiation. Fifty 15-cm plates were used for each PAR-CLIP experiment. Isolation of crosslinked RNAs, and sequencing of subsequently generated cDNA libraries were done essentially as described in ref. 28, except that wild-type and I304N FMRP isoform 1 and 7 PAR-CLIPs were performed with omission of the RNase T1 digest following the immunoprecipitation in order to increase the probability for recovering G-rich binding sites.

Recombinant baculoviral expression and purification of wild-type and I304N FMRP. Wild-type and I304N mutant versions of pENTR4-Flag-HA-FMRP-His₆ were constructed and then recombined into the pDEST 8 baculoviral expression vector by Gateway LR recombinase. Baculoviral production and amplification was performed as described in the Bac-to-Bac Manual (Invitrogen). Sf9 cells were used for the recombinant virus production, amplification and expression of recombinant FMRP proteins. Recombinant viruses were amplified to $\sim 2 \times 10^8$ plaque-forming units per ml titre. In brief, 1×10^9 Sf9 cells were infected at a multiplicity of infection (MOI) of 5 and placed back into fresh Supplemented Grace's Insect Medium (supplemented with 10% FBS, 0.1% Pluronic-F68, 100 U ml⁻¹ penicillin, 100 μ g ml⁻¹ streptomycin) into a spinner flask at a final concentration of 1×10^6 cells per ml medium for 3–4 days prior to collection. Cells were lysed in loading buffer (50 mM Tris-HCl, pH 8.0, 1 M KCl, 5 mM MgCl₂, 5 mM imidazole, 10% glycerol, 0.1% Triton X-100, 1 mM β -mercaptoethanol, 1 \times EDTA-free Protease Inhibitor Cocktail (Roche)), centrifuged (19,000g, 10 min, 4 °C) and resulting supernatants filtered through a 5- μ m membrane (Pall) prior to loading onto an XK 16 column (GE Life Science) packed with 10 ml of cobalt immobilized metal ion affinity chromatography resin (Clontech) pre-equilibrated in loading buffer and attached to an AKTA Explorer FPLC. Six column volume washes (loading buffer with 13 mM imidazole) were passed through the column and then step eluted (loading buffer with 300 mM imidazole), collecting 1-ml-sized fractions. The peak of fractions containing FMRP were determined by SDS-PAGE and Coomassie-staining of the gel. The peak fractions were pooled and dialysed overnight into buffer containing 20 mM Tris-HCl, pH 7.65, at 25 °C, 300 mM KCl, 5 mM MgCl₂, 50% glycerol, 0.1% Triton X-100, 1 mM DTT. Aliquots of FMRP were stored at -80 °C. One litre of baculoviral-infected sf9 cells typically yielded 1 mg of pure FMRP.

[γ -³²P]-ATP 5' radiolabelling of oligoribonucleotides. Synthetic oligoribonucleotides (10 pmol, 18–60-nucleotide size ranges and listed in Supplementary Tables 4 and 5) were combined with 5 pmol [γ -³²P]-ATP (6,000 Ci mmol⁻¹) in a 10- μ l reaction buffer containing 70 mM Tris-HCl, pH 7.6, at 25 °C, 10 mM MgCl₂, 5 mM DTT. The mixture was incubated at 95 °C for 30 s, and placed on ice. 5 U T4 polynucleotide kinase (NEB) was then added and incubated for 15 min at 37 °C. Thereafter, non-radiolabelled ATP was added to a final concentration of 1 mM and incubated at 37 °C for an additional 5 min at 37 °C. Following addition of 40 μ l H₂O, samples were incubated at 95 °C for 30 s, then loaded onto Microspin G25 columns (GE Life Science) and centrifuged at 730g for 1 min, according to manufacturer's instructions. The volume of the eluate was increased to 100 μ l by the addition of water for a final concentration of 100 nM and stored at -20 °C.

EMSAs. 1 nM radiolabelled RNA was combined with FMRP varying in concentration from 0–1 μ M in 20 μ l EMSA buffer (20 mM Tris-HCl, pH 7.65, 300 mM KCl, 5 mM MgCl₂, 35% glycerol, 1 mM DTT, 0.1 U μ l⁻¹ RNasin (Promega), 100 ng μ l⁻¹ transfer RNA, 0.1 mg ml⁻¹ acetylated BSA (Ambion)) and incubated at 30 °C for 1 h in 1.5 ml passivated (50 μ l 1 mg ml⁻¹ acetylated BSA, 27 °C, 1 h) and siliconized microcentrifuge tubes. After 1 h, 5 μ l of EMSA loading buffer (50% glycerol, bromophenol blue in EMSA buffer) was added. Alternatively,

for supershift assays, 1 µg antibody was added to each reaction and incubated for an additional 30 min prior to addition of EMSA loading buffer. Reactions were separated by native PAGE using 6% or 10% (bottom)/6% (top) step-gradient gels (24 × 15 cm). Polyacrylamide gels were polymerized using 49:1 acrylamide:bisacrylamide in Tris-glycine buffer (25 mM Tris base, 0.2 M Glycine), and pre-run for 30 min at 4 °C at 300 V in the same buffer. Reactions were loaded and the species were separated for 2 h at 300 V at 4 °C. The ³²P radioactive signal was detected using phosphorimager screens (1 h to overnight exposure) and the signal quantified using ImageGauge software. Curves and binding constants were calculated using Kaleidagraph v4.03 (Synergy) software.

Pre-processing reads. Details describing specific pre-processing parameters are available at <http://fmpr.rockefeller.edu>. Raw Illumina reads were stripped of 3' adapters using the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Reads shorter than 13 nucleotides or containing an ambiguous nucleotide were discarded. Processed reads were aligned to the reference genome (GRCh37/hg19) by the Bowtie algorithm (0.12.7), allowing for two alignment errors. For each read, only the best mismatch-stratum was reported for up to ten different locations; T-C mismatches with the genomic sequence were subtracted from the mismatch count for each of the mapped locations. After the conversion subtraction, reads that mapped to only one genomic location were retained for downstream analysis.

PARalyzer and annotation. For each library, PARalyzer was used to identify binding sites as described previously in ref. 34. See FMR.ini file at fmpr.rockefeller.edu for details about exact parameters. PARalyzer-derived clusters were annotated using an in-house software annotator, using annotation data from the following sources (all tables available upon request): (1) protein-coding transcript: lincRNA, misc_RNA, Mt_rRNA, Mt_tRNA, rRNA, snoRNA and snRNA annotation were downloaded from biomaRt/Ensembl version GRCh37 (<http://jun2011.archive.ensembl.org/biomart/martview>); (2) repetitive element annotation was downloaded from the RepeatMasker track of UCSC genome table browser (<http://genome.ucsc.edu/>). Repeat elements of the class ribosomal RNA, small nuclear RNA and tRNA were combined with Ensembl annotation for those categories described above; (3) piwi-interacting RNA annotation was downloaded from functional RNADB version 3.4 (http://www.ncrna.org/frnadb/files/hg18_gtf.zip) and converted to hg19 using the liftOver utility; (4) micro RNA annotation was downloaded from miRBase Sequence version 17 (<http://www.mirbase.org/>). Binding sites overlapping multiple annotation categories were assigned a single annotation according to the FMR.yaml file available at fmpr.rockefeller.edu (higher number given more priority).

cERMIT motif and RRE hierarchical analyses. We used cERMIT³⁵ to identify motifs enriched in clusters exhibiting higher T-to-C conversion evidence (for detailed parameters see cERMIT template file at fmpr.rockefeller.edu). Unlike many motif-discovery tools, cERMIT used the full data set to discover motifs with high evidence scores. We ranked binding sites mapping to 3' UTR, 5' UTR, coding or intron annotation categories by the log₂(T-to-C conversion events). Motifs found by cERMIT were filtered for low-complexity motifs (motifs for which any single nucleotide was present in 80% or more locations). To define core RREs, the position weight matrices of the remaining motifs were clustered and aligned using a modified version of the STAMP⁴⁸ tool that only considers the forward strand (available upon request).

Wild-type versus I304N comparative analysis. Because isoform 1 and 7 I304N libraries contained substantially more reads, and consequently binding sites, than

their respective wild-type libraries, we randomly sub-sampled reads from the I304N libraries such that each subset would have at least as many and no greater than ~10% more binding sites as the wild-type library. To compare RRE read depth between wild-type and I304N libraries, we first identified the genomic coordinates and underlying read depth for each instance of either ACUK or WGGA within a binding site from the wild-type library. Next, we determined the read-depth ratio comparing the wild-type RRE read depth with the read depth of the wild-type RRE coordinates in each I304N subset. We calculated the average of the I304N versus wild-type read depth ratio for each instance of an RRE in wild-type binding sites labelled as ACUK, mixed and WGGA.

RIP-chip and analyses. RIP-chip using Human Genome U133 Plus 2.0 (Affymetrix) was performed essentially as described in ref. 49. Fifty 15-cm plates of doxycycline-induced (1 µg ml⁻¹, 16–20 h) stable cells expressing Flag-HA-FMRP isoform 1 were used. Cleared supernatants were split into two equal fractions for technical replicates. 1/100 of each replicate volume was saved for total RNA isolation. The remaining lysate was incubated with magnetic beads pre-conjugated with anti-Flag M2 antibodies (10 µl conjugated beads per 1 ml of lysate). After 1-h incubation at 4 °C, beads were washed three times with IP wash buffer then re-suspended in 1 bead volume of ×1 proteinase K buffer containing 0.6 mg ml⁻¹ proteinase K, and incubated at 65 °C for 20 min. RNA was then isolated by phenol-chloroform extraction then ethanol precipitation, and re-suspended in 100 µl H₂O. RNeasy Mini Kit (Invitrogen) was used according to manufacturer's instructions and recovered RNA was ethanol-precipitated and finally dissolved in 6 µl H₂O. RNA concentration was determined using a NanoDrop UV-Vis spectrophotometer. RNA from the pre-immunoprecipitated aliquots were Trizol extracted, then similarly prepped using the RNeasy Mini Kit.

Robust multi-array average (RMA) background correction and robust quantile normalization was applied to RIP-chip microarray data using the AffyR package⁵⁰. After filtering for detection, the median of all probes mapping to a gene was used to calculate gene-level expression values. A log₂-fold enrichment (LFE) was calculated by subtracting the log₂ (RIP gene-level expression) from the log₂ (lysate gene-level expression). These LFE values were binned on the basis of properties determined by PAR-CLIP experiments and plotted in MATLAB using the empirical cumulative distribution function. For analyses involving ACUK and WGGA RREs, genes containing low-complexity motifs were excluded to determine the sole contributions of ACUK and WGGA RREs. PARalyzer-defined binding sites overlapping repetitive elements were not used for summarizing gene-level site information. The significance of enrichment for each bin of data in comparison to the total set of enrichments was determined by the Kolmogorov-Smirnov test. Bonferroni correction was applied to account for multiple comparisons.

47. Hoell, J. *et al.* in *Handbook of RNA Biochemistry* Vol. 1 (eds Hartmann, R.K., Bindereif, A., Schön, A. & Westhof, E.) (Wiley-VCH, 2013).
48. Mahony, S., Auron, P. E. & Benos, P. V. DNA familial binding profiles made easy: comparisons of various motif alignment and clustering strategies. *PLoS Comp. Biol.* **3**, e61 (2007).
49. Landthaler, M. *et al.* Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs. *RNA* **14**, 2580–2596 (2008).
50. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).

High-resolution crystal structure of human protease-activated receptor 1

Cheng Zhang¹, Yoga Srinivasan², Daniel H. Arlow³, Juan Jose Fung^{1†}, Daniel Palmer², Yaowu Zheng^{2†}, Hillary F. Green³, Anjali Pandey⁴, Ron O. Dror³, David E. Shaw³, William I. Weiss^{1,5}, Shaun R. Coughlin² & Brian K. Kobilka¹

Protease-activated receptor 1 (PAR1) is the prototypical member of a family of G-protein-coupled receptors that mediate cellular responses to thrombin and related proteases. Thrombin irreversibly activates PAR1 by cleaving the amino-terminal exodomain of the receptor, which exposes a tethered peptide ligand that binds the heptahelical bundle of the receptor to affect G-protein activation. Here we report the 2.2 Å resolution crystal structure of human PAR1 bound to vorapaxar, a PAR1 antagonist. The structure reveals an unusual mode of drug binding that explains how a small molecule binds virtually irreversibly to inhibit receptor activation by the tethered ligand of PAR1. In contrast to deep, solvent-exposed binding pockets observed in other peptide-activated G-protein-coupled receptors, the vorapaxar-binding pocket is superficial but has little surface exposed to the aqueous solvent. Protease-activated receptors are important targets for drug development. The structure reported here will aid the development of improved PAR1 antagonists and the discovery of antagonists to other members of this receptor family.

Protease-activated receptors (PARs) are G-protein-coupled receptors (GPCRs) that mediate cellular responses to specific proteases^{1,2}. The coagulation protease thrombin activates the prototypical PAR, PAR1, by specific cleavage of the N-terminal exodomain of the receptor to generate a new N terminus. This new N terminus then functions as a tethered peptide agonist that binds intramolecularly to the seven-transmembrane helix bundle of the receptor to affect G-protein activation^{1,3–8} (Fig. 1a). In adult mammals, the four members of the PAR family link tissue injury and local generation of active coagulation proteases to cellular responses that help to orchestrate haemostasis, thrombosis, inflammation and perhaps tissue repair^{2,9}. PARs may also participate in the progression of specific cancers^{10,11}.

In contrast to a typical receptor–agonist binding interaction, the interaction of PAR1 with its activator, thrombin, is that of a protease substrate, with thrombin binding transiently to the receptor, cleaving it, then dissociating^{1,3–12}. Proteolytic unmasking of the tethered peptide agonist of the receptor is irreversible, and although a free synthetic hexapeptide with the amino acid sequence of the tethered agonist (SFLLRN) can activate the receptor with half-maximum effective concentration (EC₅₀) values in the 3–10-μM range, the local concentration of the tethered agonist peptide is estimated to be about 0.4 mM. Accordingly, PAR signalling must be actively terminated^{13–15} and, unlike most other GPCRs that can go through many rounds of activation by reversible diffusible hormones and neurotransmitters, PARs are degraded after a single activation^{6,13–17}. Identification of effective PAR antagonists has been challenging because low molecular mass compounds must compete with the very high local concentration of the tethered agonist generated by proteolytic cleavage.

Vorapaxar is a highly specific, virtually irreversible PAR1 antagonist¹⁸ (Supplementary Fig. 1). In a phase III trial, vorapaxar protected patients against recurrent myocardial infarction at a cost of increased bleeding^{19,20}. Given the latter, an antagonist that is reversible in the setting of bleeding might be desirable. Although the very slow dissociation rate of vorapaxar from PAR1 probably accounts for its ability to

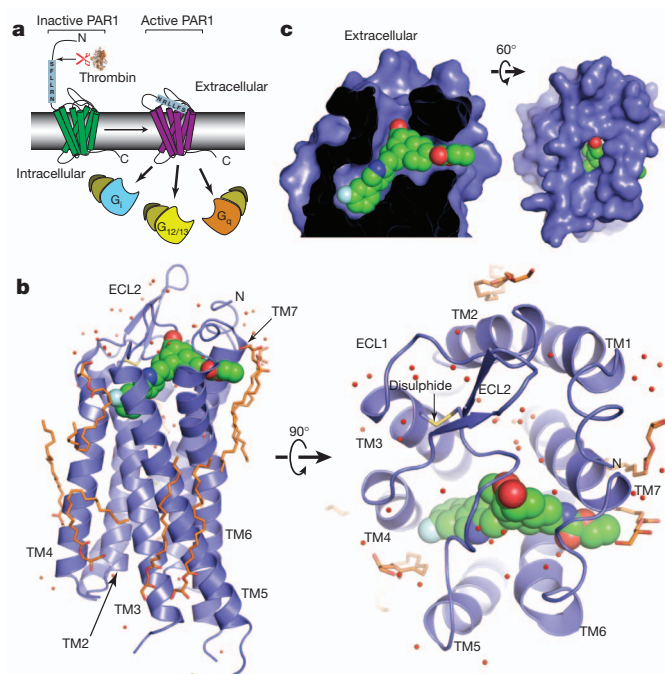


Figure 1 | PAR1 activation and overall structure of human PAR1 complex with antagonist vorapaxar. a, Thrombin cleaves the PAR1 N terminus and exposes a new N-terminal peptide, SFLLRN, which can bind to and activate the transmembrane core of PAR1. PAR1 can activate several G proteins including G_i, G_{12/13} and G_q. **b**, Overall view of the human PAR1 structure and the extracellular surface. The receptor is shown as blue ribbon and vorapaxar is shown as green spheres. Monoolein is shown in orange, water in red. The disulphide bond is shown as a yellow stick. **c**, Surface view of the ligand-binding pocket viewed from two different perspectives. The vorapaxar-binding pocket is close to the extracellular surface but not well exposed to the extracellular solvent.

¹Department of Molecular and Cellular Physiology, Stanford University School of Medicine, Stanford, California 94305, USA. ²Cardiovascular Research Institute, University of California, San Francisco, 555 Mission Bay Boulevard South, S452P, San Francisco, California 94158, USA. ³D. E. Shaw Research, New York, New York 10036, USA. ⁴Portola Pharmaceuticals, 270 East Grand Avenue, South San Francisco, California 94080, USA. ⁵Department of Structural Biology, Stanford University School of Medicine, 299 Campus Drive, Stanford, California 94305, USA. [†]Present addresses: ProNovus Bioscience, LLC, 544 E Weddell Drive, Sunnyvale, California 94089, USA (J.J.F.); Northeast Normal University, Changchun, Jilin 130024, China (Y.Z.).

inhibit receptor activation by its tethered agonist peptide, it may be possible to develop a drug with an off rate slow enough to block signalling but fast enough to allow useful reversal after cessation of drug.

In an effort to advance our understanding of PAR1 structure and function and to provide a foundation for discovery of new agents to advance the pharmacology of PARs, we obtained the crystal structure of vorapaxar-bound human PAR1.

Crystallization of human PAR1

To facilitate crystallogenes, T4 lysozyme (T4L) was inserted into intracellular loop 3 (ICL3) in human PAR1, the N-linked glycosylation sites in ECL2 were mutated²¹, and the N-terminal exodomain was removed by site-specific cleavage at a tobacco etch virus (TEV) protease site introduced between amino acids 85 and 86 (ref. 4; Supplementary Fig. 2). The structure of human PAR1–T4L bound to vorapaxar was determined to 2.2 Å by merging diffraction data sets from 18 crystals grown in the lipidic cubic phase (Supplementary Figs 3 and 4). Details of data collection and structure refinement are listed in Supplementary Table 1.

PAR1 has the expected seven-transmembrane segment bundle (Fig. 1b). There are several lipid molecules assigned as monoolein from lipidic cubic phase in the structure (Fig. 1b), but no ordered cholesterol molecules were observed. The remaining N-terminal fragment Arg 86–Glu 90 and a part of the ICL2 from Gln 209 to Trp 213 are not modelled in the structure because of the weak electron density. There is no clear electron density for residues after Cys 378, and no helix 8 is observed after transmembrane segment 7 (TM7) in the structure. Whether this reflects a lack of a helix 8 in PAR1 in its native state or conditions in the crystal is not known.

Cys 175^{3,25} in helix III and Cys 254 in extracellular loop 2 (ECL2) form a conserved disulphide bond (Figs 1b and 2a). Amino-terminal to Cys 254, ECL2 loops outwards in two anti-parallel β strands. This structural feature is found in other peptide receptors, including the CXCR4 receptor and the opioid receptors^{22–25}, despite absence of amino acid sequence homology among these receptors in ECL2 (Supplementary Fig. 5). In contrast to the open, solvent-exposed binding pocket observed in the μ -opioid receptor (MOR) and other peptide receptors, access to the vorapaxar-binding pocket is restricted by the central location of ECL2 (Fig. 1 and Supplementary Figs 5 and 6), which almost completely covers the extracellular-facing surface of vorapaxar. ECL2 is anchored in this position by hydrogen bonds between His 255 in ECL2 and Tyr 353^{7,35} in TM7, and between Asp 256 in ECL2 and Tyr 95 in the N terminus (Supplementary Fig. 6), and by extensive interactions with vorapaxar (Fig. 2). The covered vorapaxar-binding pocket in PAR1 more closely resembles rhodopsin and the lipid-activated sphingosine-1-phosphate receptor (S1P₁) than other peptide-activated GPCRs (Supplementary Fig. 5).

Divergence of PAR1 from other family A GPCRs

Members of the family A GPCRs share a set of conserved amino acids that are thought to be important in signal transduction^{26,27}. Specific residues in the highly conserved FXXCWXP motif (in which X denotes any amino acid) in TM6, and the NPXXY motif in TM7 undergo structural rearrangements during activation of the β_2 adrenergic receptor (β_2 AR) (Fig. 3b, d). However, based on a phylogenetic analysis of amino acid sequences²⁸, PAR1 is a more distant relative of the family A GPCRs that have been crystallized thus far. PAR1 belongs to the δ -subfamily, which includes the glycoprotein receptors, the purinergic receptors and the olfactory receptors²⁸. The tryptophan residue in FXXCW^{6,48}XP proposed to act as a toggle switch during activation in some GPCRs²⁸ is replaced by Phe^{6,48} in all PARs (Fig. 3a). Phe^{6,44}, also highly conserved in family A GPCRs, is Phe^{6,44} in PAR1, but Tyr^{6,44} in PAR2 and Ala^{6,44} in PAR4. When comparing inactive and active states of the β_2 AR, changes in packing interactions involving Pro^{5,50}, Ile^{3,40} and Phe^{6,44} seem to have a role in structural changes needed to accommodate G-protein binding^{29–31}. Packing

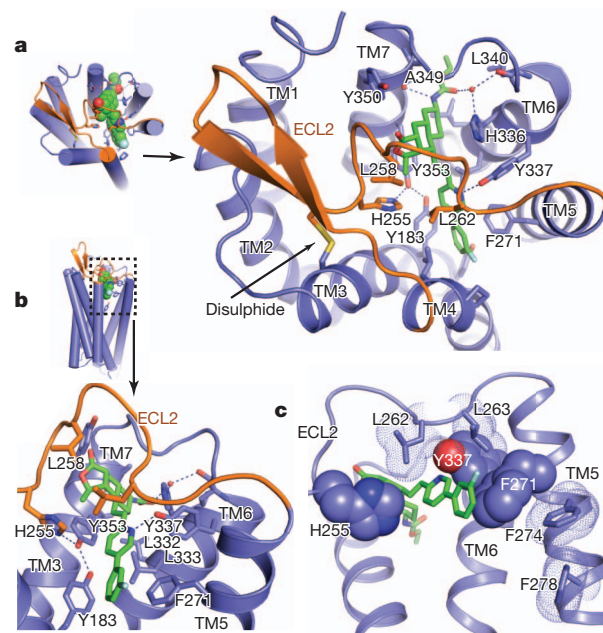


Figure 2 | Binding interactions of vorapaxar with human PAR1.

a, b, Ligand-binding pocket viewed from the extracellular surface (**a**) and from the side of the transmembrane helix bundle (**b**). ECL2 is coloured in orange in **a** and **b**. Ligand vorapaxar is shown as green sticks. Water molecules are shown as red spheres. Hydrogen bonds are shown as dotted lines. **c**, Two residues, Leu 262 and Leu 263 in ECL2 (shown as dot surface), which pack against residues His 255, Phe 271^{5,39} and Tyr 337^{6,59} (shown in Corey–Pauling–Koltun (CPK) representation), may contribute to the selectivity of vorapaxar for human PAR1. Also shown are Phe 274^{5,42} and Phe 278^{5,46} in TM5 (shown as dot surface), which may indirectly influence vorapaxar binding by packing interactions with Phe 271^{5,39}.

interactions of the corresponding residues Pro 282^{5,50}, Ile 190^{3,40} and Phe 322^{6,44} in the PAR1 differ from those in both active and inactive β_2 AR structures (Fig. 3b). Taken together, these differences suggest that PAR1 may differ from other family A GPCRs in the mechanism by which signals propagate from the extracellular peptide-binding interface to the cytoplasmic domains that interact with G proteins and other signalling molecules.

The NP^{7,50}XXY motif at the end of helix VII observed in most family A GPCRs is DP^{7,50}XXY in PAR1. This region undergoes structural rearrangement after activation of the β_2 AR. In PAR1, Asp 367^{7,49} and Tyr 371^{7,53} form hydrogen bonds with residues in TM2 and TM1 (Fig. 3c). The hydrogen-bonding network associated with Asp 367^{7,49} is extensive and includes several water molecules and a putative sodium ion. Na⁺, rather than a water molecule, was assigned to this region of electron density as it has five oxygen neighbours and short distances to its oxygen ligands (average refined distance 2.4 Å), both consistent with known Na⁺–oxygen interactions³², and it interacts with two acidic side chains that, assuming deprotonated states, would repel one another without charge neutralization provided by the Na⁺. The sodium ion also interacts with a conserved Asp 148^{2,50} in TM2 and Ser 189^{3,39} in TM3, with two water molecules nearby (Fig. 3c). Na⁺ is an allosteric modulator for several family A GPCRs such as the α_2 A adrenergic receptor, A_{2A} adenosine receptor, μ - and δ -opioid receptors and D₂ dopamine receptor^{33–37}. The conserved Asp^{2,50} is necessary for sodium sensitivity of the α_2 A adrenergic receptor³⁷ and D₂ dopamine receptor^{36,38}. In PAR1, Asp 367^{7,49} might be expected to form a stronger hydrogen-bonding network and sodium coordination site than asparagine residues found in most other family A GPCRs. This more stable network may contribute to the unusual position of the cytoplasmic end of TM7 that is displaced inward towards TM2. This position is more similar to the

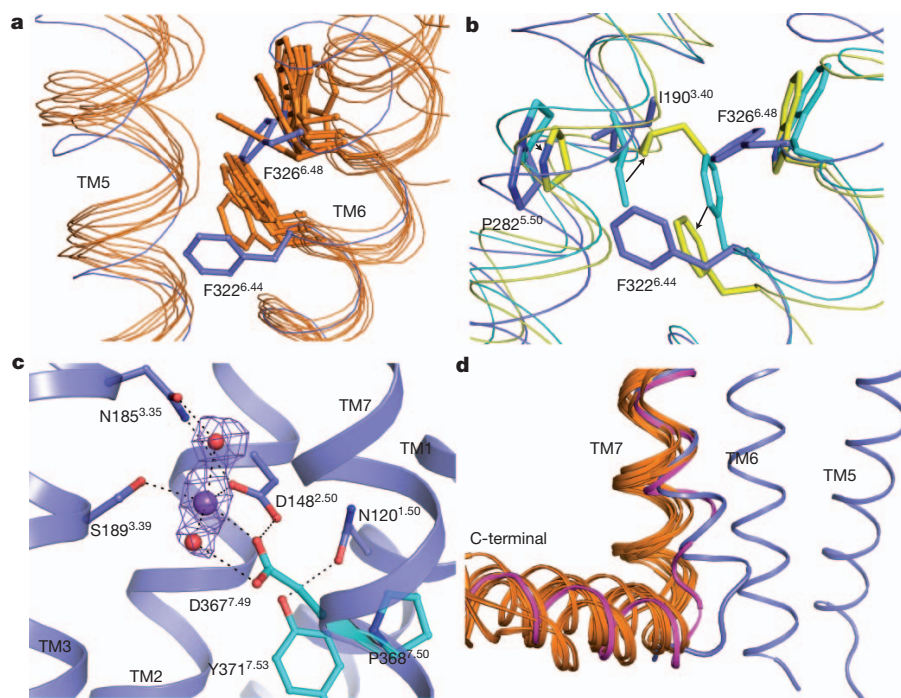


Figure 3 | Structure motifs in PAR1 compared with other family A GPCRs.

a, Superimposition of TM5 and TM6 of human PAR1 (in blue) with those of other GPCRs including β_2 AR and β_1 AR, A_{2A} adenosine receptor, dopamine D_3 receptor, M2 muscarinic receptor, histamine H1 receptor, μ -opioid receptor, S1P₁ and CXCR4 (all in orange). Phe 326^{6.48} and Phe 322^{6.44} in the F^{6.44}XXC(F)^{6.48}XP motif in PAR1 are shown as sticks. This motif is FXXC(W)^{6.48}XP in most other family A GPCRs. Phe 326^{6.48} and Phe 322^{6.44} are both in different conformations compared to their counterparts in other GPCRs. **b**, In the β_2 AR, rearrangements of three residues, Pro^{5.50}, Ile^{3.40} and Phe^{6.44}, are associated with receptor activation. Black arrows indicate changes of these residues going from inactive (cyan) to active (yellow) β_2 AR structures. The counterparts in the inactive state structure of PAR1 (Pro 282^{5.50}, Ile 190^{3.40}

and Phe 322^{6.44}) are shown in blue. **c**, DP^{7.50}XXY motif in TM7 and sodium-binding site in PAR1. Residues Asp 367^{7.49}, Pro 368^{7.50} and Tyr 371^{7.53} in the DP^{7.50}XXY motif are shown as cyan sticks. This motif is normally NPXXY in most other family A GPCRs. Sodium is shown as a purple sphere and water molecules are shown as red spheres. Polar interactions are shown as black dashed lines. An $F_o - F_c$ omit electron-density map for the putative sodium ion and water molecules contoured at 4σ is shown as purple mesh.

d, Superimposition of the C-terminal part of TM7 in the structure of human PAR1 (blue), in the inactive structures of other GPCRs (all in orange) mentioned in **a** and in the active structure of β_2 AR (in magenta). The C-terminal part of TM7 in PAR1 adopts a conformation more similar to that observed in the active state of the β_2 AR.

active β_2 AR bound with either nanobody 80 or heterotrimeric G protein^{29,31} (Fig. 3d).

Structural insights into binding properties of vorapaxar

Vorapaxar binds in an unusual location very close to the extracellular surface of PAR1. By contrast, ligands for other GPCRs penetrate more deeply into the transmembrane core (Fig. 1 and Supplementary Figs 5 and 7). The vorapaxar-binding pocket, composed of residues from TM3, TM4, TM5, TM6 and TM7 as well as ECL2 and ECL3, forms a tunnel across the receptor with one end open between TM4 and TM5 and the other between TM6 and TM7 occupied by the ethyl carbamate tail of vorapaxar (Figs 1 and 2). There is only a small opening in the extracellular surface between ECL2 and ECL3. Details of interactions between vorapaxar and PAR1 are illustrated in Fig. 2 and Supplementary Fig. 8.

Vorapaxar shows high selectivity for human PAR1 over human PAR2 and PAR4, and mouse PAR1 in functional assays (Supplementary Fig. 9A, B). The structural basis for this selectivity is not readily apparent from the crystal structure. Nearly all the residues that interact with vorapaxar in human PAR1 are conserved in human PAR2, human PAR4 and mouse PAR1 (Supplementary Fig. 10). Residues Leu 262 and Leu 263, which are involved in weak hydrophobic interactions with vorapaxar in human PAR1, are alanine and asparagine, respectively, in human PAR4, and Leu 263 is a methionine in mouse PAR1 (Fig. 2c and Supplementary Fig. 10). These differences by themselves would not be expected to explain the high selectivity of vorapaxar. However, Leu 262 and Leu 263 pack

against other amino acids that have more extensive interactions with vorapaxar. Leu 262 interacts with His 255 in ECL2 and Leu 263 interacts with Phe 271^{5.39} at the top of TM5 and Tyr 337^{6.59} at the top of TM6 (Fig. 2c). These interactions may influence ligand-binding selectivity indirectly by contributing to the overall structure and stability of the binding pocket. Amino acid differences between PAR1, PAR2 and PAR4 more distant from the ligand-binding pocket may also contribute to subtype-specific binding of vorapaxar. Phe 274^{5.42} is Leu in PAR2 and PAR4, and Phe 278^{5.46} is Val in PAR2 and Gly in PAR4. Although neither Phe 274^{5.42} nor Phe 278^{5.46} directly contact vorapaxar, Phe 278^{5.46} packs against Phe 274^{5.42}, which in turn packs against Phe 271^{5.39} in the binding pocket (Fig. 2c).

In human PAR2 and PAR4, ECL3 connecting TM6 and TM7 is one residue shorter than it is in PAR1 (Supplementary Fig. 9C). Tyr 337^{6.59} at the carboxy-terminal end of TM6 forms a strong hydrogen bond with vorapaxar (Fig. 2 and Supplementary Fig. 9D). Another residue, Tyr 353^{7.35}, which forms the base of the ligand-binding pocket together with Tyr 183^{3.33}, is located at the N-terminal end of TM7 (Fig. 2 and Supplementary Fig. 9D). A shorter ECL3 in human PAR2 and PAR4 may change the relative position of these amino acids, thereby altering the overall geometry of the binding pocket. Although the length of ECL3 in mouse and human PAR1 is the same, four of the eight amino acids are different (Supplementary Fig. 9C). These differences may affect the structure of ECL3 and thereby influence interactions between vorapaxar in Tyr 337^{6.59} and Tyr 353^{7.35}. Alternatively, these differences could have an effect on the mechanism by which vorapaxar gains access to the binding pocket.

Structural insights into vorapaxar inhibition of PAR1

Although this structure is compatible with the very slow dissociation rate of vorapaxar, it does not provide insight into the mechanism by which vorapaxar or an agonist peptide gains access to the binding pocket. None of the three openings to the vorapaxar-binding pocket is large enough to accommodate the passage of the ligand. We thus wondered whether the unliganded receptor might have a more open structure, similar to that observed for opioid receptors, with unique interactions between vorapaxar and PAR1 causing an otherwise open binding pocket to close after vorapaxar binding. To investigate this issue, we performed long-timescale molecular dynamics simulations of PAR1 with and without vorapaxar bound. Intriguingly, removal of the ligand did not lead to a more open binding pocket, but instead to one that was even more closed (Fig. 4a, b, Supplementary Fig. 11 and Supplementary Table 2). The extracellular end of TM6 moved about 4 Å inward towards TM4, bringing ECL3 in full contact with ECL2 and completely occluding the binding pocket. By contrast, in a similar study on the MOR, the binding pocket remained open when the ligand was removed (Supplementary Fig. 11). The collapse of the vorapaxar-binding pocket may reflect the fact that both vorapaxar and its binding pocket are uncharged, whereas in the opioid receptors and many other family A GPCRs, the charged residue Asp^{3.32} helps to keep the binding pocket hydrated after the ligand is removed.

It is interesting to speculate that vorapaxar, a highly lipophilic molecule, may access the binding pocket through the lipid bilayer, possibly between TM6 and TM7. This is similar to the binding mode proposed for retinal to rhodopsin and the lipid sphingosine-1-phosphate (S1P) to the S1P₁ receptor^{39,40} (Supplementary Fig. 5).

To understand the ability of vorapaxar to inhibit agonist binding and activation, we examined the functional consequences of mutating four aromatic amino acids that form strong interactions with vorapaxar: Tyr 183^{3.33}, Tyr 353^{7.35}, Phe 271^{5.39} and Tyr 337^{6.59} (Fig. 2a, b and Supplementary Fig. 3C). Three of these (Tyr 353^{7.35}, Phe 271^{5.39} and Tyr 337^{6.59}) assume substantially different positions in simulations of unliganded PAR1 (Fig. 4a). Tyrosine residues Tyr 183^{3.33} in TM3 and Tyr 353^{7.35} in TM7 are linked to each other by a hydrogen bond and form the base of the binding 'tunnel', and are part of a

hydrophobic cage that surrounds the ligand. Tyr 353^{7.35} also forms a hydrogen bond with His 255, the most deeply buried amino acid in ECL2 (Fig. 2a and Supplementary Fig. 6). This interaction of His 255 with Tyr 353^{7.35} contributes to the closed conformation of ECL2 over the ligand-binding pocket. Phe 271^{5.39} interacts with the fluoro-phenyl ring and Tyr 337^{6.59} forms a strong hydrogen bond with the pyridine ring of vorapaxar.

Mutation of Tyr 337^{6.59} to phenylalanine and Tyr 353^{7.35} to alanine led to a reduction in cell surface expression, making it difficult to interpret the associated reduction in agonist peptide activation (Fig. 4c, d). Mutation of Phe 271^{5.39} to alanine was associated with enhanced cell surface expression, but reduced activation by agonist peptide. There was little effect of this mutation on maximal inhibition by vorapaxar. Although not conclusive, this result suggests that Phe 271^{5.39} may have a role in both peptide and vorapaxar binding. Of interest, mutation of Tyr 183^{3.33} exhibited enhanced response to the agonist peptide and loss of inhibition by vorapaxar (Fig. 4c). This result suggests a possible role for Tyr 183^{3.33} in maintaining the receptor in an inactive state and indicates that interactions between Tyr 183^{3.33} and vorapaxar may further stabilize an inactive conformation.

Activation of PAR1 by the agonist peptide

Our structure is consistent with data from mutagenesis studies that suggest that the PAR1-tethered agonist peptide may activate the hepta-helical bundle of the receptor by interacting with superficial structures rather than penetrating deeply into the transmembrane core^{8,41–45}. Glu 260, a solvent-exposed residue in ECL2 in both vorapaxar-bound and unliganded PAR1 (Fig. 5), is of particular interest, as evidence from mutagenesis studies suggests an interaction with Arg 46 in the tethered peptide SFLLRN^{8,44}. Substitution of Glu 260 with arginine markedly reduces activation of PAR1 by a peptide with the native tethered ligand sequence (SFLLRN) but facilitates activation by SFLLEN. Arginine substitution of Glu 264, which is surface-exposed and near Glu 260 in the structure, also facilitates activation by SFLLEN.

Mutation of other residues near the extracellular surface, including Leu96Ala (N terminus), Asp256Ala (ECL2) and Glu347^{7.29} Ala/Gln (ref. 45) (Fig. 5a), markedly reduces activation of PAR1 by the peptide

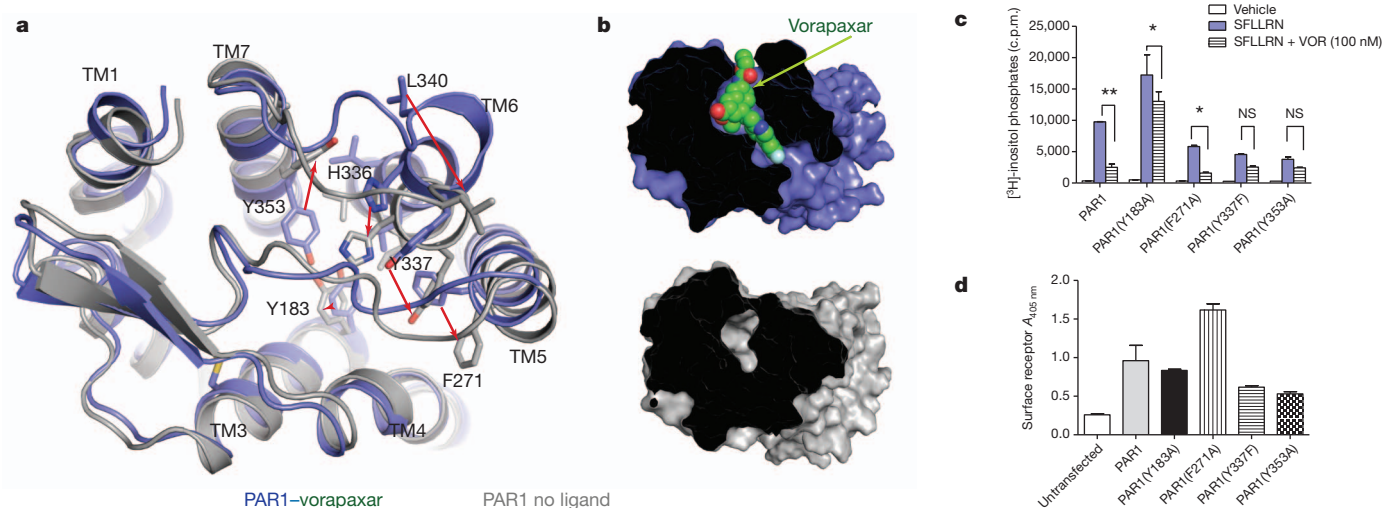


Figure 4 | Collapse of ligand-binding pocket in long-timescale molecular dynamics simulations of unliganded PAR1. Molecular dynamics simulations were performed on PAR1 from which vorapaxar had been removed. The vorapaxar-bound PAR1 crystal structure is shown in blue and the unliganded structure obtained from molecular dynamics simulation is shown in grey. **a**, The largest differences between vorapaxar-bound and unliganded PAR1 are at the extracellular end of TM6 and in ECL3. Residues involved in vorapaxar binding are shown as sticks. **b**, Surface view showing collapse of the ligand-binding pocket during molecular dynamics simulation in the absence of vorapaxar. **c**, **d**, Signalling (**c**) and cell surface (**d**) expression for wild-type

human PAR1 and binding-site mutants. Cos7 cells expressing the indicated receptor constructs were labelled with [³H]-myo-inositol, pretreated with vehicle or 100 nM vorapaxar in DMEM medium containing 0.1% BSA, 20 mM HEPES and 0.2% β-hydroxy cyclodextrin (to retain vorapaxar in solution) for 1 h, then incubated with vehicle or PAR1 agonist (100 μM SFLLRN) for 1 h at 37 °C. Total [³H]-inositol-phosphate accumulation was measured. Surface expression of receptors in cells transfected in parallel was assessed by measuring binding of anti-Flag antibody to an epitope displayed at the N terminus of the receptor. Error bars, mean ± s.e. Results are representative of three separate experiments. NS, not significant.

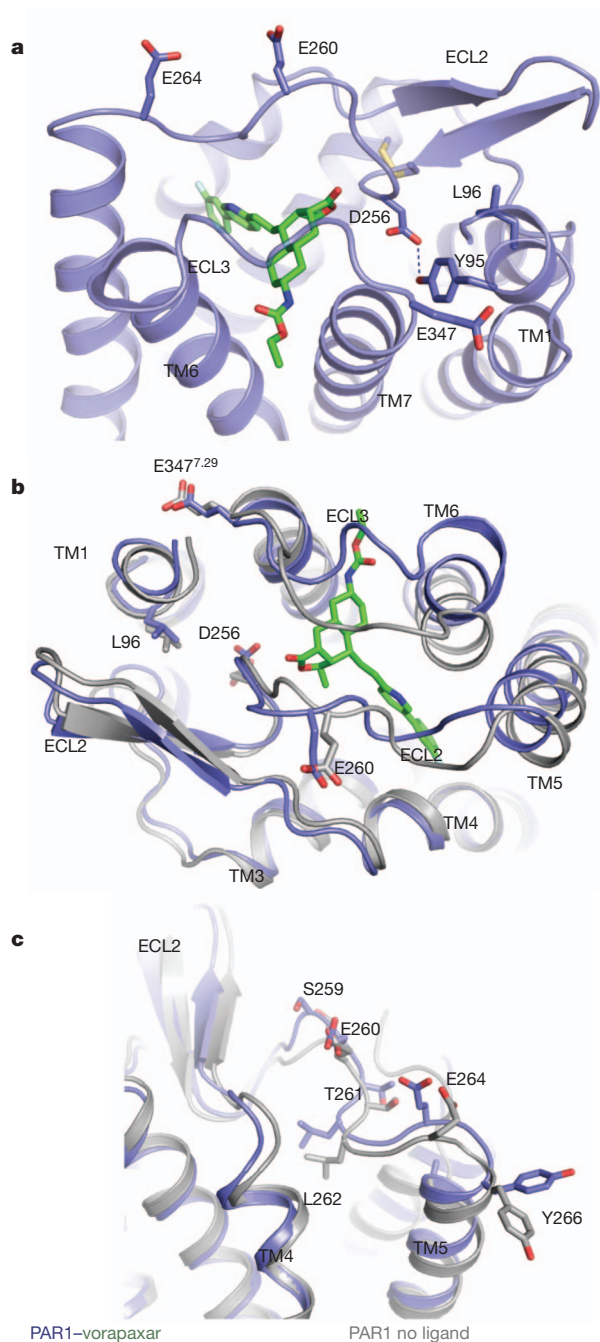


Figure 5 | Residues important for agonist peptide binding and receptor activation. **a**, Mutations of residues Glu 260, Asp 256, Leu 96 and Glu 347^{7,29} near the extracellular surface have been shown to reduce activation of PAR1 by the free agonist peptide. Among them only Glu 260 is completely exposed to the solvent, whereas Asp 256 is the most deeply buried, forming a hydrogen bond with residue Tyr 95. Although none of these amino acids forms part of the vorapaxar-binding pocket, Asp 256 forms a hydrogen bond with Tyr 95 that may stabilize ECL2 over the vorapaxar-binding pocket. Vorapaxar is shown in green. **b, c**, Superimposition of the unliganded molecular dynamics simulation model (grey) with the ligand-bound crystal structure (blue). In **b**, Residues Glu 260, Asp 256, Leu 96 and Glu 347^{7,29}, which are important for agonist peptide signalling, are in similar positions in both structures. **c**, The positions of residues that differ between human and *Xenopus* PAR1 in ECL2. Substitution of these residues in human PAR1 with corresponding residues from *Xenopus* PAR1 results in increased basal activity.

agonist. However, these mutations have only a small effect on agonist peptide binding, with only Asp256Ala resulting in a more than tenfold loss of binding affinity⁴³. The positions of these residues do not change

substantially when comparing vorapaxar-bound and unliganded PAR1 (Fig. 5b). Of interest, only Glu 347^{7,29} is surface exposed (Fig. 5a, b), suggesting that Leu 96 and Asp 256 may not interact directly with the agonist peptide or that these amino acids are more exposed than would appear from the molecular dynamics model of the unliganded receptor. In the inactive structure, Asp 256 forms a hydrogen bond with Tyr 95 and helps to stabilize interactions between the C-terminal end of the N terminus and ECL2 (Fig. 5a).

Interestingly, substitution of human PAR1 sequence Asn 259–Ala 268, the region of ECL2 implicated in tethered ligand binding, with the cognate *Xenopus* ECL2 sequence results in an approximately tenfold increase in basal activity⁴⁶. Figure 5c shows the position of amino acids that differ between human and *Xenopus* receptors in ECL2 in both the crystal structure and the unliganded molecular dynamics simulation. The superficial location of these activating mutations suggests that very superficial interactions between the tethered agonist peptide and the extracellular loops may be sufficient to activate PAR1. Taken together, these mutagenesis studies suggest that the PAR1 agonist peptide may activate PAR1 by binding more superficially than do agonist peptides for opioid receptors. Alternatively, the tethered peptide may bind in a sequential manner, initially to the extracellular loops but penetrating more deeply into the core of the receptor through a sequence of conformational intermediates.

Conclusion

The unusual mode of activation and the paucity of pharmacological tools have made PAR1 one of the more challenging GPCRs to characterize and a difficult target for drug development. The crystal structure offers insights into the very high affinity interaction with the antagonist vorapaxar. This structure will provide a template for the development of PAR1 antagonists with better drug properties and the development of antagonists for other PAR subtypes to probe their biological roles. The mechanism of activation of PAR1 remains poorly understood. Molecular dynamics simulations of an unliganded receptor together with the location of amino acids known to influence agonist peptide activity suggest that activation of PAR1 by its agonist peptide may involve superficial interactions with extracellular loops. Future efforts will focus on an active-state structure of PAR1 bound to its tethered agonist peptide.

METHODS SUMMARY

The human PAR1–T4L fusion protein was expressed in Sf9 insect cells and purified by nickel-affinity chromatography, Flag M1 antibody affinity chromatography followed by size exclusion chromatography. PAR1–T4L crystals were grown using the *in meso* crystallization method. The diffraction data were collected from 18 crystals at the GM/CA@APS beamline in the Argonne National Laboratory. The structure was solved by molecular replacement and refined in Phenix. Refinement statistics are provided in Supplementary Table 1. All-atom molecular dynamics simulations were performed on Anton⁴⁷ with lipids and water molecules represented explicitly. Phosphoinositide hydrolysis assays were done in Cos7 cells transfected with wild-type and mutant PAR1. More details are provided in Methods.

Full Methods and any associated references are available in the online version of the paper.

Received 6 August; accepted 22 October 2012.

Published online 9 December 2012.

1. Vu, T. K., Hung, D. T., Wheaton, V. I. & Coughlin, S. R. Molecular cloning of a functional thrombin receptor reveals a novel proteolytic mechanism of receptor activation. *Cell* **64**, 1057–1068 (1991).
2. Coughlin, S. R. Thrombin signalling and protease-activated receptors. *Nature* **407**, 258–264 (2000).
3. Vu, T. K., Wheaton, V. I., Hung, D. T., Charo, I. & Coughlin, S. R. Domains specifying thrombin-receptor interaction. *Nature* **353**, 674–677 (1991).
4. Chen, J., Ishii, M., Wang, L., Ishii, K. & Coughlin, S. R. Thrombin receptor activation. Confirmation of the intramolecular tethered ligand hypothesis and discovery of an alternative intermolecular liganding mode. *J. Biol. Chem.* **269**, 16041–16045 (1994).

5. Liu, L. W., Vu, T. K., Esmon, C. T. & Coughlin, S. R. The region of the thrombin receptor resembling hirudin binds to thrombin and alters enzyme specificity. *J. Biol. Chem.* **266**, 16977–16980 (1991).
6. Ishii, K., Hein, L., Kobilka, B. & Coughlin, S. R. Kinetics of thrombin receptor cleavage on intact cells. Relation to signaling. *J. Biol. Chem.* **268**, 9780–9786 (1993).
7. Mathews, I. I. *et al.* Crystallographic structures of thrombin complexed with thrombin receptor peptides: existence of expected and novel binding modes. *Biochemistry* **33**, 3266–3279 (1994).
8. Gerszten, R. E. *et al.* Specificity of the thrombin receptor for agonist peptide is defined by its extracellular surface. *Nature* **368**, 648–651 (1994).
9. Coughlin, S. R. Protease-activated receptors in hemostasis, thrombosis and vascular biology. *J. Thromb. Haemost.* **3**, 1800–1814 (2005).
10. García-López, M. T., Gutierrez-Rodríguez, M. & Herranz, R. Thrombin-activated receptors: promising targets for cancer therapy? *Curr. Med. Chem.* **17**, 109–128 (2010).
11. Ramachandran, R., Noorbakhsh, F., Defea, K. & Hollenberg, M. D. Targeting proteinase-activated receptors: therapeutic potential and challenges. *Nature Rev. Drug Discov.* **11**, 69–86 (2012).
12. Ishii, K. *et al.* Determinants of thrombin receptor cleavage. Receptor domains involved, specificity, and role of the P3 aspartate. *J. Biol. Chem.* **270**, 16435–16440 (1995).
13. Shapiro, M. J., Trejo, J., Zeng, D. & Coughlin, S. R. Role of the thrombin receptor's cytoplasmic tail in intracellular trafficking. Distinct determinants for agonist-triggered versus tonic internalization and intracellular localization. *J. Biol. Chem.* **271**, 32874–32880 (1996).
14. Trejo, J. & Coughlin, S. R. The cytoplasmic tails of protease-activated receptor-1 and substance P receptor specify sorting to lysosomes versus recycling. *J. Biol. Chem.* **274**, 2216–2224 (1999).
15. Trejo, J., Hammes, S. R. & Coughlin, S. R. Termination of signaling by protease-activated receptor-1 is linked to lysosomal sorting. *Proc. Natl Acad. Sci. USA* **95**, 13698–13702 (1998).
16. Shapiro, M. J. & Coughlin, S. R. Separate signals for agonist-independent and agonist-triggered trafficking of protease-activated receptor 1. *J. Biol. Chem.* **273**, 29009–29014 (1998).
17. Hein, L., Ishii, K., Coughlin, S. R. & Kobilka, B. K. Intracellular targeting and trafficking of thrombin receptors. A novel mechanism for resensitization of a G protein-coupled receptor. *J. Biol. Chem.* **269**, 27719–27726 (1994).
18. Chackalamannil, S. *et al.* Discovery of potent orally active thrombin receptor (protease activated receptor 1) antagonists as novel antithrombotic agents. *J. Med. Chem.* **48**, 5884–5887 (2005).
19. Morrow, D. A. *et al.* Vorapaxar in the secondary prevention of atherothrombotic events. *N. Engl. J. Med.* **366**, 1404–1413 (2012).
20. Kosoglou, T. *et al.* Pharmacodynamics and pharmacokinetics of the novel PAR-1 antagonist vorapaxar (formerly SCH 530348) in healthy subjects. *Eur. J. Clin. Pharmacol.* **68**, 249–258 (2012).
21. Soto, A. G. & Trejo, J. N-linked glycosylation of protease-activated receptor-1 second extracellular loop: a critical determinant for ligand-induced receptor activation and internalization. *J. Biol. Chem.* **285**, 18781–18793 (2010).
22. Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071 (2010).
23. Manglik, A. *et al.* Crystal structure of the micro-opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321–326 (2012).
24. Wu, H. *et al.* Structure of the human κ -opioid receptor in complex with JDTic. *Nature* **485**, 327–332 (2012).
25. Granier, S. *et al.* Structure of the δ -opioid receptor bound to naltrindole. *Nature* **485**, 400–404 (2012).
26. Smit, M. J. *et al.* Pharmacogenomic and structural analysis of constitutive G protein-coupled receptor activity. *Annu. Rev. Pharmacol. Toxicol.* **47**, 53–87 (2007).
27. Schwartz, T. W., Frimurer, T. M., Holst, B., Rosenkilde, M. M. & Elling, C. E. Molecular mechanism of 7TM receptor activation—a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* **46**, 481–519 (2006).
28. Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol. Pharmacol.* **63**, 1256–1272 (2003).
29. Rasmussen, S. G. *et al.* Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature* **469**, 175–180 (2011).
30. Rosenbaum, D. M. *et al.* Structure and function of an irreversible agonist- β_2 adrenoceptor complex. *Nature* **469**, 236–240 (2011).
31. Rasmussen, S. G. *et al.* Crystal structure of the β_2 adrenergic receptor–Gs protein complex. *Nature* **477**, 549–555 (2011).
32. Harding, M. M. Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. *Acta Crystallogr. D* **58**, 872–874 (2002).
33. Horstman, D. A. *et al.* An aspartate conserved among G-protein receptors confers allosteric regulation of α_2 -adrenergic receptors by sodium. *J. Biol. Chem.* **265**, 21590–21595 (1990).
34. Costa, T., Lang, J., Gless, C. & Herz, A. Spontaneous association between opioid receptors and GTP-binding regulatory proteins in native membranes: specific regulation by antagonists and sodium ions. *Mol. Pharmacol.* **37**, 383–394 (1990).
35. Gao, Z. G. & Ijzerman, A. P. Allosteric modulation of A_{2A} adenosine receptors by amiloride analogues and sodium ions. *Biochem. Pharmacol.* **60**, 669–676 (2000).
36. Selent, J., Sanz, F., Pastor, M. & De Fabritiis, G. Induced effects of sodium ions on dopaminergic G-protein coupled receptors. *PLOS Comput. Biol.* **6**, e1000884 (2010).
37. Wilson, M. H., Highfield, H. A. & Limbird, L. E. The role of a conserved inter-transmembrane domain interface in regulating α_{2A} -adrenergic receptor conformational stability and cell-surface turnover. *Mol. Pharmacol.* **59**, 929–938 (2001).
38. Neve, K. A. *et al.* Modeling and mutational analysis of a putative sodium-binding pocket on the dopamine D2 receptor. *Mol. Pharmacol.* **60**, 373–381 (2001).
39. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289**, 739–745 (2000).
40. Hanson, M. A. *et al.* Crystal structure of a lipid G protein-coupled receptor. *Science* **335**, 851–855 (2012).
41. Lerner, D. J., Chen, M., Tram, T. & Coughlin, S. R. Agonist recognition by proteinase-activated receptor 2 and thrombin receptor. Importance of extracellular loop interactions for receptor function. *J. Biol. Chem.* **271**, 13943–13947 (1996).
42. Nanevich, T. *et al.* Mechanisms of thrombin receptor agonist specificity. Chimeric receptors and complementary mutations identify an agonist recognition site. *J. Biol. Chem.* **270**, 21619–21625 (1995).
43. Blackhart, B. D. *et al.* Extracellular mutations of protease-activated receptor-1 result in differential activation by thrombin and thrombin receptor agonist peptide. *Mol. Pharmacol.* **58**, 1178–1187 (2000).
44. Seeley, S. *et al.* Structural basis for thrombin activation of a protease-activated receptor: inhibition of intramolecular liganding. *Chem. Biol.* **10**, 1033–1041 (2003).
45. Bahou, W. F., Kutok, J. L., Wong, A., Potter, C. L. & Collier, B. S. Identification of a novel thrombin receptor sequence required for activation-dependent responses. *Blood* **84**, 4195–4202 (1994).
46. Nanevich, T., Wang, L., Chen, M., Ishii, M. & Coughlin, S. R. Thrombin receptor activating mutations. Alteration of an extracellular agonist recognition domain causes constitutive signaling. *J. Biol. Chem.* **271**, 702–706 (1996).
47. Shaw, D. E. *et al.* Millisecond-scale molecular dynamics simulations on Anton. *Proc. Conf. High Performance Computing Networking, Storage Analysis* Article no. 39 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge support from the National Institutes of Health, grants NS028471 (B.K.K.), and HL44907 and HL65590 (S.R.C.), and from the Mathers Foundation (B.K.K. and W.I.W.).

Author Contributions C.Z. optimized the construct, expressed and purified human PAR1-T4L for crystallization, developed the purification procedure, performed crystallization trials, optimized crystallization conditions, collected diffraction data, and solved and refined the structure. Y.S. helped design and make constructs for baculoviral expression, and executed and analysed cell-based functional assays of wild-type and mutant PAR1. D.H.A. designed, performed and analysed molecular dynamics simulations. J.J.F. developed the initial expression and purification protocol for PAR1. D.P. helped design, execute and analyse platelet function studies. Y.Z. helped design and make constructs for baculoviral PAR expression. H.F.G. performed and analysed molecular dynamics simulations, and assisted with manuscript preparation. A.P. made vorapaxar. R.O.D. oversaw, designed and analysed molecular dynamics simulations, and assisted with manuscript preparation. D.E.S. oversaw molecular dynamics simulations and analysis. W.I.W. assisted with X-ray diffraction data processing and crystal structure refinement. S.R.C. and B.K.K. initiated the project, planned and analysed experiments, and supervised the research and wrote the manuscript with C.Z.

Author Information Atomic coordinates and structure factors for PAR1 are deposited in the Protein Data Bank under accession code 3VW7. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.R.C. (shaun.coughlin@ucsf.edu) or B.K.K. (kobilka@stanford.edu).

METHODS

PAR1–T4L expression and purification. To facilitate crystallization, a human PAR1 construct was generated with several modifications. A TEV protease-recognition site was introduced after residue Pro 85, two N-linked glycosylation sites in ECL2 were removed by mutation (Asn250Gly and Asn259Ser), and the C terminus was truncated after residue Ser 395. T4L residues 2–161 (ref. 48) were inserted into the third intracellular loop between residues Ala 301 and Ala 303, with only one residue Val 302 removed. To facilitate purification, an N-terminal Flag epitope was inserted after a signal peptide and a C-terminal deca-histidine tag was introduced. The final crystallization construct PAR1–T4L is shown in Supplementary Fig. 2.

The modified PAR1 was expressed in Sf9 cells using the pFastBac baculovirus system (Invitrogen). The ligand vorapaxar was added at 100 nM to the cells during expression. The cells were infected with baculovirus at 27 °C for 48 h before collection. To purify the receptor, infected cells were lysed by osmotic shock in low-salt buffer containing 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 100 mM vorapaxar and 2 mg ml⁻¹ iodoacetamide. Iodoacetamide was used to alkylate reactive cysteines to prevent nonspecific oligomerization. The protein was further extracted from cell membranes using a glass dounce homogenizer in buffer containing 20 mM Tris-HCl, pH 7.5, 350 mM NaCl, 1% dodecyl maltoside (DDM), 0.03% cholesterol hemisuccinate (CHS), 0.2% sodium cholate, 10% glycerol, 2 mg ml⁻¹ iodoacetamide and 100 nM vorapaxar. Cell debris was removed by high-speed centrifugation. From this point, 1 µM vorapaxar was added in all the following buffers used for purification. Nickel-NTA agarose resin was added to the supernatant after homogenization and stirred for 1 h at 4 °C. The resin was then washed three times in batch with buffer comprised of 20 mM HEPES, pH 7.5, 350 mM NaCl, 0.1% DDM, 0.02% CHS and 1 µM vorapaxar, and transferred to a glass column. The bound receptor was eluted with buffer containing 300 mM imidazole and loaded onto an anti-Flag M1 affinity column. After extensive washing with buffer comprised of 20 mM HEPES, pH 7.5, 350 mM NaCl, 0.1% DDM, 0.02% CHS, 1 µM vorapaxar and 2 mM Ca²⁺, the receptor was eluted from M1 resin using the same buffer without Ca²⁺ but with 200 µg ml⁻¹ Flag peptide and 5 mM EDTA. To remove extra N-terminal residues and the Flag epitope, TEV protease was added to the receptor and the cleavage reaction run at room temperature overnight. Size exclusion chromatography was used to obtain the final monodisperse receptor preparation. Purified PAR1–T4L was concentrated to 40–50 mg ml⁻¹ using 100 kDa cut-off Vivaspinn concentrators for crystallization.

Crystallization. As for other T4L-fused GPCRs crystallized so far, *in meso* crystallization was used to obtain PAR1–T4L crystals^{49,50}. The protein was mixed with monolein and cholesterol (10:1 by mass) using the two syringe mixing method by weight of 1:1.5 (protein:lipid). After a clear lipidic cubic phase formed, the mix was dispensed onto glass plates in 20–40 nl drops overlaid with 700 nl precipitant solution using a Gryphon LCP robot. Crystals appeared in two days in 0.1–0.2 M sodium chloride, 100 mM sodium phosphate, pH 6.0–6.5, 25–35% PEG300, and grew to full size after 1 week (Supplementary Fig. 3).

Data collection and structure determination. Crystals were collected and frozen in liquid nitrogen. Data collection was performed at beamline 23-ID of GM/CA@APS at the Advanced Photon Source. Microbeams of 10 or 20 µm diameter were used to acquire all diffraction data. Owing to radiation damage, only 5–20 degrees of rotation data were collected from each crystal. All data were processed with the HKL2000 package⁵¹. A 2.2-Å data set was obtained by merging diffraction data from 18 crystals. The space group was determined to be *P*₂₁₂₁₂₁. Molecular replacement was performed using the program Phaser⁵² in Phenix⁵³, with the CXCR4 structure (PDB accession 3ODU) as the search model. The seven-transmembrane helices without any loops, and the T4L in the CXCR4 structure, were used as independent search models. The initial structure model was completed and improved through iterative refinement in Phenix⁵³ and manual rebuilding of all the loops and several parts in the transmembrane region in Coot⁵⁴. Model refinement in Phenix and manual adjustment in Coot was performed to improve the model. The final structure was determined at 2.2-Å resolution. The quality of the structure was assessed using Molprobity⁵⁵. Data processing and structure refinement statistics are shown in Supplementary Table 1. Figures were prepared with PyMol⁵⁶.

Phosphoinositide hydrolysis assays and cell-surface expression level. The QuikChange site-directed mutagenesis kit (Agilent) was used to generate human PAR1 mutants and all mutants were fully sequenced. Cos7 cells were transiently transfected using Eugene HD with either empty vector or wild-type human PAR1 and mutants in the mammalian expression vector pBJ1 and signalling assays were performed as described in ref. 46. In brief, Cos7 cells expressing wild-type or mutant human PAR1 were labelled with [³H]-myo-inositol, then incubated with vehicle or 100 nM vorapaxar in DMEM medium containing 0.1% BSA, 20 mM HEPES, 0.2% 2-hydroxypropyl-β-cyclodextrin (to retain vorapaxar in solution)

for 1 h at 37 °C. Agonist (100 µM SFLLRN or 10 nM thrombin for PAR1 or other PAR agonists as indicated) was added and incubation continued for 1 h. The total amount of accumulated [³H]-inositol phosphates accumulated was determined as in ref. 46. Cos7 cells transfected with empty vector had little response to PAR agonists, and treatment with vorapaxar alone did not affect phosphoinositide hydrolysis (Supplementary Fig. 9A).

Surface expression of receptors was measured as described in refs 13 and 16. Cos7 cells were transiently transfected with empty pBJ1 or pBJ1 directing expression of N-terminal Flag-tagged versions of wild-type human PAR1 or mutants. After 48 h, cells were washed once with serum-free medium containing 0.1% BSA and 20 mM HEPES, then incubated with 3 µg ml⁻¹ Flag M1 antibody (Sigma) for 1 h at 4 °C in the same medium. The cells were then washed twice with PBS containing Ca²⁺ and Mg²⁺ to remove unbound antibody and fixed with 2% paraformaldehyde for 5 min. The cells were then washed twice with PBS with Ca²⁺ and Mg²⁺, and incubated with goat anti-mouse horseradish peroxidase (HRP)-conjugated secondary antibody, washed and developed with one-step ABTS HRP substrate (Pierce). The absorbance at 405 nm was measured as indication of cell surface receptor expression levels.

Platelet signalling assays. Washed human platelets were prepared and PAR1-dependent responses were measured as described in ref. 57. In brief, acid-citrate-dextrose anti-coagulated human blood samples (60 ml per donor) were obtained from AllCells. Blood was centrifuged without braking at 250g at 37 °C for 15 min. The upper platelet-rich plasma phase was collected, incubated at 37 °C for 10 min in the presence of prostacyclin (PGI₂, 0.5 µM), and centrifuged at 2,200g for 15 min. The pellet was resuspended in complete Tyrode's solution (134 mM NaCl, 12 mM NaHCO₃, 2.9 mM KCl, 0.34 mM Na₂HPO₄, 1.0 mM MgCl₂, 10 mM HEPES, 0.9% (w/v) dextrose, pH 7.4) containing 2 mM CaCl₂, 0.35% (w/v) bovine serum albumin (BSA), 10 U ml⁻¹ heparin, and 0.5 µM PGI₂. The platelet suspension was incubated for 10 min at 37 °C then centrifuged at 1,900g for 8 min. This wash step was repeated and the final pellet resuspended in Tyrode's buffer supplemented with BSA and 0.02 U ml⁻¹ apyrase. Platelets were incubated at 37 °C for 30 min to allow recovery from the effects of PGI₂, counted using a Hemavet FS950 (Drew Scientific) and diluted to 300,000 cells per micro-litre in Tyrode's solution.

To antagonize PAR1, vorapaxar or vehicle (2% (w/v) 2-hydroxypropyl-β-cyclodextrin in dimethylsulphoxide (DMSO)) were added to platelet suspensions that were then incubated for 1 h at 37 °C before addition of agonists. The final concentrations of 2-hydroxypropyl-β-cyclodextrin and DMSO in platelet suspensions were 0.002% and 0.1%, respectively. Where reversibility was evaluated, platelets were washed twice with Tyrode's buffer containing BSA and PGI₂ after vorapaxar treatment then diluted for cell activation assays as above.

For flow cytometric analysis of platelet activation, platelets suspended in Tyrode's solution containing 2 mM CaCl₂, 0.35% BSA and 0.02 U ml⁻¹ apyrase were incubated with agonist and antibody simultaneously. Fluor-conjugated antibodies directed against P-selectin (phycoerythrin (PE)-conjugate of AK-4; Ebiosciences; 1:25 dilution) and the activated conformation of integrin α_{IIb}β₃ (FITC-conjugate of PAC-1, BD Biosciences; 1:25 dilution) were used. After 15 min at 37 °C, the platelet suspension was diluted with PBS and platelet-bound antibody measured using an Accuri C6 flow cytometer (Accuri). Samples from at least three different donors were analysed, each in triplicate.

Molecular dynamics simulation methods. In all simulations, the receptor was embedded in a hydrated lipid bilayer with all atoms, including those in the lipids and water, represented explicitly. Simulations were performed on Anton⁴⁷, a special-purpose computer designed to accelerate molecular dynamics simulations by orders of magnitude.

System set-up and simulation protocol. Simulations of PAR1 were based on the crystal structure of the PAR1–vorapaxar complex. The crystallized construct has T4L inserted into ICL3 in place of residue 302. For the simulations, the T4L portion was omitted, and residue 302 was modelled in. The unresolved segment of ICL2 (residues 209–213) was also modelled in. Residues 209 and 213 were added manually, and residues 210–212 were modelled in using Prime (Schrödinger LLC). The Refine Loops tool in Prime, with default settings, was then used to refine residues 209–213.

The simulation of the MOR dimer was based on the crystal structure of MOR bound to the irreversible antagonist β-funaltrexamine (PDB accession 4DKL). Both monomers of the crystallographic dimer were included in the simulation, but β-funaltrexamine was deleted from the binding pocket. As with PAR1, the T4L sequence was omitted in our simulations. Side chains for residue Met 65^{1,29}, Thr 67^{1,31}, Lys 260^{ICL3} and Arg 263^{ICL3} were not fully resolved in the crystal structure, so they were modelled in by hand, with rotamers chosen to avoid any clashes with resolved residues.

For both PAR1 and MOR, hydrogens were added to the crystal structures using Maestro (Schrödinger LLC), as described in previous work⁵⁸. Histidines were

singly protonated on the epsilon nitrogen. All other titratable residues were left in their dominant protonation state for pH 7.0, except for Asp 367^{7,49} in PAR1 and Asp 114^{2,50} in MOR, which were protonated, and Asp 148^{2,50} in PAR1, which was protonated in certain simulations. PAR1 was simulated both with and without the crystallographic sodium ion by Asp 148^{2,50} (Supplementary Table 2); Asp 148^{2,50} was not protonated in simulations that included this ion, but was protonated otherwise. The conserved aspartate at position 2.50 is known to be protonated in rhodopsin³⁹, and the residue at position 7.49 is most often an (uncharged) asparagine residue in family A GPCRs (the 'N' of the NPXXY motif).

Prepared protein structures were inserted into an equilibrated POPC bilayer as described previously³². Sodium and chloride ions were added to neutralize the net charge of the system and to create a 150-mM solution.

Simulations of the PAR1 receptor initially measured $88.9 \times 88.9 \times 88.7$ Å and contained 174 lipid molecules, and approximately 13,152 water molecules, for a total of ~67,500 atoms. When the crystallographic sodium ion near Asp 148^{2,50} was included, the simulation contained 32 sodium ions and 36 chloride ions. When the crystallographic sodium ion was not included, the system contained 31 sodium ions and 36 chloride ions. To simulate the unliganded PAR1 receptor, vorapaxar was deleted from the binding pocket. Simulations of the MOR dimer initially measured $100.0 \times 100.0 \times 89.0$ Å and contained 204 lipid molecules, 19 sodium ions, 43 chloride ions, and approximately 16,654 water molecules, for a total of ~86,700 atoms.

All simulations were equilibrated using Anton in the NPT ensemble at 310 K (37 °C) and 10^5 Pa with $5 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ harmonic position restraints applied to all non-hydrogen atoms of the protein and the ligand; these restraints were tapered off linearly over 50 ns. All bond lengths to hydrogen atoms were constrained using M-SHAKE⁶⁰. A RESPA integrator⁶¹ was used with a time step of 2 fs, and long-range electrostatics were computed every 6 fs. Production simulations were initiated from the final snapshot of the corresponding equilibration runs, with velocities sampled from the Boltzmann distribution at 310 K, using the same integration scheme, long-range electrostatics method, temperature and pressure. For PAR1, Van der Waals and short-range electrostatic interactions were cut-off at 10.3 Å and long-range electrostatic interactions were computed using the *k*-space Gaussian split Ewald method⁶² with a $32 \times 32 \times 32$ grid, $\sigma = 2.27$ Å, and $\sigma_s = 1.59$ Å. For MOR, van der Waals and short-range electrostatic interactions were cut-off at 10.16 Å and long-range electrostatic interactions were computed using the *k*-space Gaussian split Ewald method with a $64 \times 64 \times 64$ grid, $\sigma = 2.25$ Å, and $\sigma_s = 1.55$ Å.

We performed two vorapaxar-bound PAR1 simulations and four unliganded PAR1 simulations, and results were consistent across each set. The two receptors in our MOR dimer simulation also exhibited consistent behaviour. The simulation protocol we followed has been validated in previous simulations of GPCRs^{63,64}. Nevertheless, it is possible that different behaviour might have been observed in even longer simulations, with different force field parameters, or with a different choice of simulation conditions.

Force field parameters. We used the CHARMM27 parameter set for protein molecules and salt ions, with the CHARMM TIP3P water model⁶⁵; protein parameters incorporated CMAP terms⁶⁶ and modified charges on the Asp, Glu and Arg side chains⁶⁷. We used a modified CHARMM lipid force field⁶⁸. Force field parameters for vorapaxar were obtained from the CHARMM ParamChem web server⁶⁹, version 0.9.6 β.

Analysis protocols. Trajectory snapshots, each containing a record of all atom positions at a particular instant in time, were saved every 180 ps during

production simulations. Time series data shown in Supplementary Fig. 11 were smoothed by applying a 9.9-ns (55-snapshot) running average. VMD was used to visualize trajectories⁷⁰.

48. Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into β_2 -adrenergic receptor function. *Science* **318**, 1266–1273 (2007).
49. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature Protocols* **4**, 706–731 (2009).
50. Caffrey, M. Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu. Rev. Biophys.* **38**, 29–51 (2009).
51. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
52. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
53. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
54. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
55. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
56. The PyMOL Molecular Graphics System v.1.5.0.4 (Schrödinger, LLC, 2010).
57. Cornelissen, I. *et al.* Roles and interactions among protease-activated receptors and P2ry12 in hemostasis and thrombosis. *Proc. Natl Acad. Sci. USA* **107**, 18605–18610 (2010).
58. Dror, R. O. *et al.* Identification of two distinct inactive conformations of the β_2 -adrenergic receptor reconciles structural and biochemical observations. *Proc. Natl Acad. Sci. USA* **106**, 4689–4694 (2009).
59. Fahmy, K. *et al.* Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. *Proc. Natl Acad. Sci. USA* **90**, 10206–10210 (1993).
60. Krätzer, V., Gunsteren, W. F. v. & Hünenberger, P. H. A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J. Comput. Chem.* **22**, 501–508 (2001).
61. Tuckerman, M., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97**, 1990–2001 (1992).
62. Shan, Y., Klepeis, J. L., Eastwood, M. P., Dror, R. O. & Shaw, D. E. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. *J. Chem. Phys.* **122**, 54101 (2005).
63. Dror, R. O. *et al.* Activation mechanism of the β_2 -adrenergic receptor. *Proc. Natl Acad. Sci. USA* **108**, 18684–18689 (2011).
64. Dror, R. O. *et al.* Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl Acad. Sci. USA* **108**, 13118–13123 (2011).
65. MacKerell, A. D. Jr, Bashford, D. & Bellott, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).
66. Mackerell, A. D. Jr, Feig, M. & Brooks, C. L. III. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **25**, 1400–1415 (2004).
67. Piana, S., Lindorff-Larsen, K. & Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **100**, L47–L49 (2011).
68. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).
69. Vanommeslaeghe, K. *et al.* CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **31**, 671–690 (2010).
70. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

Dynamical age differences among coeval star clusters as revealed by blue stragglers

F. R. Ferraro¹, B. Lanzoni¹, E. Dalessandro¹, G. Beccari², M. Pasquato¹, P. Miocchi¹, R. T. Rood^{3,†}, S. Sigurdsson⁴, A. Sills⁵, E. Vesperini⁶, M. Mapelli⁷, R. Contreras¹, N. Sanna¹ & A. Mucciarelli¹

Globular star clusters that formed at the same cosmic time may have evolved rather differently from the dynamical point of view (because that evolution depends on the internal environment) through a variety of processes that tend progressively to segregate stars more massive than the average towards the cluster centre¹. Therefore clusters with the same chronological age may have reached quite different stages of their dynamical history (that is, they may have different ‘dynamical ages’). Blue straggler stars have masses greater² than those at the turn-off point on the main sequence and therefore must be the result of either a collision^{3,4} or a mass-transfer event^{5–7}. Because they are among the most massive and luminous objects in old clusters, they can be used as test particles with which to probe dynamical evolution. Here we report that globular clusters can be grouped into a few distinct families on the basis of the radial distribution of blue stragglers. This grouping corresponds well to an effective ranking of the dynamical stage reached by stellar systems, thereby permitting a direct measure of the cluster dynamical age purely from observed properties.

We have analysed the entire database of blue straggler stars (BSSs) collected by our group for a sample of 21 globular clusters (see Supplementary Information). Such a data set contains clusters with nearly the same chronological age (12–13 Gyr (ref. 8); the only exception is Palomar 14, which formed ~ 10.5 Gyr ago⁹) but with very different structural properties (and hence possibly at different stages of dynamical evolution). Although significant variations in the radial distribution of BSSs between clusters are already known^{10,11}, we have found that, when the radial distance is expressed in units of the core radius (to permit a meaningful comparison between the clusters), the BSS distributions seem surprisingly similar within distinct subsamples. These similarities are so striking that clusters can be efficiently grouped on the basis of the shape of their BSS radial distribution, and at least three distinct families can be defined. The observational panorama is summarized in Figs 1–3, in which the BSS distribution is compared with that of a reference population (typically red giants or horizontal-branch stars; see Supplementary Information).

Preliminary results^{12,13} have shown that the observed radial distribution of BSSs is primarily modelled by the long-term effect of dynamical friction acting on the cluster binary population (and its progeny) since the early stages of cluster evolution. In fact, whereas BSSs generated by stellar collisions are expected to be the main or sole contributors to the central peak of the distribution¹⁴, the portion beyond the cluster core, where the minimum of the distribution is observed, is entirely due to BSSs generated by mass transfer or merger in primordial binary systems, in agreement with what is found to be the dominant formation channel in other low-density environments such as open clusters¹⁵. In particular, what we call mass-transfer BSSs today are the by-product of the evolution of a $\sim 1.2M_{\odot}$ primordial binary that has been orbiting the cluster and suffering the effects of dynamical

friction for a significant fraction of the cluster’s lifetime. Hence, the radial distribution of BSSs that is now observed simply reflects the underlying distribution of $1.2M_{\odot}$ binaries, which has been shaped by dynamical friction for several billion years (see Supplementary Information).

Dynamical friction has the effect of driving objects that are more massive than the average towards the cluster centre, with an efficiency that decreases for increasing radial distance as a function of the velocity dispersion and mass density^{13,16}. Hence, as time passes, heavy objects

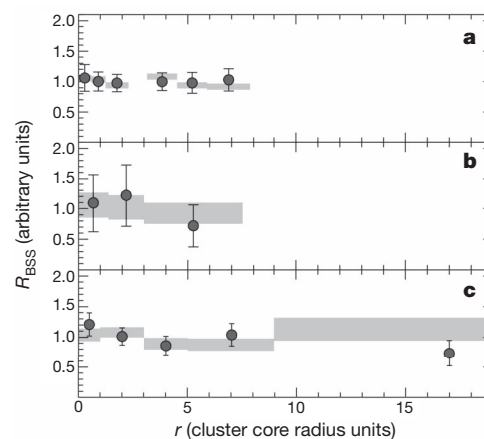


Figure 1 | The radial distribution of BSSs in three dynamically young stellar systems (family I). **a**, ω Centauri; **b**, Palomar 14; **c**, NGC 2419. The double-normalized ratio of BSSs (R_{BSS} ; dots) is defined¹⁰ as $R_{\text{BSS}}(r) = [N_{\text{BSS}}(r)/N_{\text{BSS,tot}}]/[L_{\text{BSS}}(r)/L_{\text{BSS,tot}}]$, where $N_{\text{BSS}}(r)$ is the number of BSSs measured in any given radial bin, $N_{\text{BSS,tot}}$ is the total number of such stars, and $L_{\text{BSS}}(r)$ and $L_{\text{BSS,tot}}$ are the analogous quantities for the sampled luminosity. Grey regions correspond to the double-normalized ratio measured for the reference population (red giants or horizontal-branch stars). Error bars and the width of the grey bands (1σ) have been computed from the error propagation law, by assuming Poissonian number counts and a few per cent uncertainty in the fraction of sampled luminosity, respectively. For a meaningful cluster-to-cluster comparison, the distance from the centre (r) is expressed in units of the cluster core radius. Simple theoretical arguments²⁵ demonstrate that the double-normalized ratio is equal to unity for any population (such as red giants and horizontal-branch stars) whose radial distribution follows that of the cluster’s integrated luminosity. In the three cases plotted here, BSSs show no evidence of mass segregation with respect to the reference population at any distance from the centre (note that essentially the entire radial extension is sampled by the observations). This is the most direct evidence that these stellar systems are dynamically unevolved, with mass segregation not yet being established even in the central regions. Our conclusions are further strengthened by the fact that ω Centauri is not now considered to be a genuine globular cluster²⁶ but instead the remnant of a dwarf galaxy; in fact, no signs of mass segregation are expected in collisionless systems.

¹Department of Physics and Astronomy, University of Bologna, Viale Berti Pichat 6/2, 40127 Bologna, Italy. ²European Southern Observatory, Karl Schwarzschild Strasse 2, D-85748 Garching bei München, Germany. ³Astronomy Department, University of Virginia, PO Box 400325, Charlottesville, Virginia 22904, USA. ⁴Department of Astronomy and Astrophysics, Pennsylvania State University, 525 Davey Laboratory, University Park, Pennsylvania 16802, USA. ⁵Department of Physics and Astronomy, McMaster University, Hamilton, Ontario L8S 4M1, Canada. ⁶Department of Astronomy, Indiana University, Bloomington, Indiana 47405, USA. ⁷INAF–Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, 35122 Padova, Italy.

†Deceased.

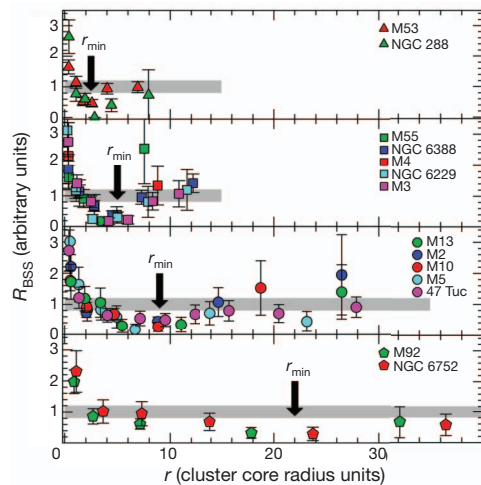


Figure 2 | The radial distribution of BSSs in systems of intermediate dynamical ages (family II). For the sake of clarity, the grey strips schematically represent the reference population distributions (which are shown in Supplementary Fig. 1 and in specific papers describing each individual cluster^{10–13,17,20,24}). The radial distributions of BSSs (large coloured symbols, 1σ errors) are clearly incompatible with that of the reference populations: they appear bimodal, with a well-defined peak in the cluster centre (testifying to a strong central segregation), a dip at intermediate radii (r_{\min} ; see Supplementary Information) and a rising branch in the outskirts. Clusters have been grouped according to the value of r_{\min} (thick arrows): from top to bottom, the minimum is observed at progressively larger distances from the centre. This radius marks the distance at which dynamical friction has already been effective in segregating BSSs towards the cluster centre. Hence, in contrast with those plotted in Fig. 1, these systems show evidence of dynamical evolution, progressively increasing from top to bottom. According to this interpretation, M53 and NGC 288 should be the dynamically youngest of the clusters of intermediate dynamical age. Note that, in spite of its possible appearance, there is no correlation between the extent of the observations and the value of r_{\min} ; in a few cases the most external point is not plotted (M53, 47 Tuc and M3) for the sake of clarity. Moreover, as a result of insufficient quality of data or strong contamination by Galactic field stars, the most external part of the radial distribution of BSSs is lacking in a few clusters (NGC 6388, M4 and NGC 6229). However, r_{\min} is well detected in all cases and these drawbacks do not affect the conclusion of the paper.

orbiting at larger and larger distances from the cluster centre are expected to drift towards the core and their radial distribution to develop a peak in the cluster centre and a dip (that is, a region devoid of these stars) that progressively propagates outwards. As the dynamical evolution of the system proceeds, the portion of the cluster where dynamical friction has been effective increases and the radial position of the minimum of the distribution (r_{\min} ; see Supplementary Information) increases. In spite of the crude approximations, even a simple analytical estimate¹⁶ of the radius at which dynamical friction is expected to segregate $1.2M_{\odot}$ stars over the lifetime of the cluster has been found to be in excellent agreement with the observed value of r_{\min} in a few globular clusters^{13,17}. The progressive outward drift of r_{\min} as a function of time is fully confirmed by the results that we obtained from direct N -body simulations that followed the evolution of $1.2M_{\odot}$ objects within a ‘reference’ cluster over a significant fraction of its lifetime (see Supplementary Information).

In view of these considerations, the families defined in Figs 1–3 correspond to clusters of increasing dynamical ages. The signature of the parent cluster’s dynamical evolution encoded in the BSS population has now been finally deciphered: the shape of the radial distribution of BSSs is a powerful indicator of dynamical age. A flat radial distribution of BSSs (consistent with that of the reference population, as found for family I in Fig. 1) indicates that dynamical friction has not yet had a major effect even in the innermost regions, and the cluster is still dynamically young. This situation is confirmed by observations of dwarf

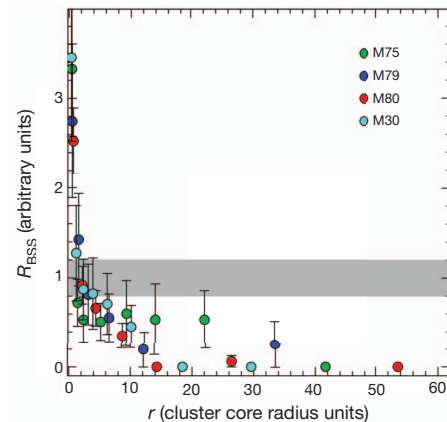


Figure 3 | The radial distribution of BSSs in dynamically old clusters (family III). The grey strip is as in Fig. 2. The BSS radial distributions in this family of clusters are monotonic, with only a central peak followed by a rapid decline and no signs of an external rising branch; these systems therefore show the highest level of dynamical evolution, with even the farthest BSSs already sunk towards the cluster centre. Even if in this regime the dynamical clock were to start to saturate, a ranking could still be attempted on the basis of the shape of the BSS distribution: M75 (green dots), where some BSSs are still orbiting at $r \approx 20$ and the slope of the decreasing branch is flatter, could be the dynamically youngest cluster within the family; M80 (red dots), with very sharply decreasing distribution, could have the highest dynamical age. Because our observations sample almost the entire radial extension of each cluster, we are confident that no BSS rising branch is present beyond the limit reached by the data. Error bars indicate 1σ .

spheroidal galaxies: for these collisionless systems we do not expect dynamical friction to be efficient, and indeed no statistically significant dip in the distribution of BSSs has been observed^{18,19}. In more evolved clusters (family II in Fig. 2), dynamical friction starts to be effective and segregates heavy objects that are orbiting at distances still relatively close to the centre; as a consequence, a peak in the centre and a minimum at small radii appear in the BSS distribution. Meanwhile, the most remote BSSs have not yet been affected by the action of dynamical friction (this generates the rising branch of the observed bimodal BSS distributions). Because the action of dynamical friction extends progressively to larger and larger distances from the centre, the dip of the distribution moves progressively outwards (as seen in the different groups of family II clusters). In highly evolved systems we expect that even the most remote BSSs were affected by dynamical friction and started to drift gradually towards the centre. As a consequence the external rising branch of the radial distribution disappears (as observed for family III in Fig. 3). All the clusters showing BSS distribution with only a central peak can therefore be classified as ‘dynamically old’. This class includes M30, a system that has already experienced core collapse^{20,21}, which is considered to be a typical symptom of extreme dynamical evolution¹ (see Supplementary Information).

The proposed classification is also able to shed light on several controversial cases that have been debated in the literature, thus further demonstrating the importance of a reliable determination of the cluster’s dynamical age. In fact, in contrast with previous studies²² suggesting that the core of M4 might have collapsed, we find that M4 belongs to a family of clusters of intermediate dynamical age. NGC 6752 turns out to be in a relatively advanced state of dynamical evolution, possibly on the verge of core collapse, as also suggested by its double King profile indicating that the cluster core is detaching from the rest of the cluster structure²³. Finally, this approach might provide the means of discriminating between a central density cusp due to core collapse (as for M30)²⁰ and that due to the presence of an exceptional concentration of dark massive objects (neutron stars and/or the long-sought and still elusive intermediate-mass black holes; see the case of NGC 6388 (ref. 24)).

The quantization into distinct age-families is of course an oversimplification: in nature a continuous behaviour is expected and the

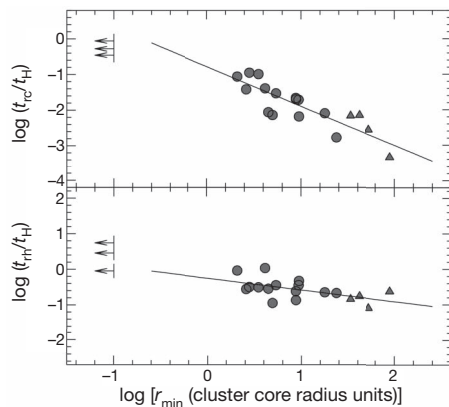


Figure 4 | A first calibration of the clock. The relaxation times at the cluster centre (t_{rc}) and at the half-mass radius (t_{rh}), normalized to the age of the Universe ($t_H = 13.7$ Gyr), are plotted as a function of the hand of our clock (r_{min} , in units of the cluster core radius). Relaxation times have been computed by following the literature²⁷, using accurately re-determined values of the structural parameters (derived from the King-model²⁸ fit to the observed star density profiles^{11,17,20,24}) and a homogeneous distance scale²⁹. The dynamically young systems (family I), showing no minimum, are plotted as lower-limit arrows at $r_{min} = 0.1$ and are not used to derive the best-fit relations (solid lines). For dynamically old clusters (family III, triangles) we adopted $r_{min} = r_0$, where r_0 is the distance from the centre of the most distant bin where no BSSs are found. As expected for a meaningful clock, a tight anticorrelation is found: clusters with relaxation times of the order of the age of the Universe show no signs of BSS segregation (hence the radial distribution of BSSs is flat and r_{min} is not definable; see Fig. 1), whereas for decreasing relaxation times the radial position of the minimum increases progressively.

position of r_{min} should vary with continuity as a sort of clock hand. This allows us to push our analysis further and define the first empirical clock able to measure the dynamical age of a stellar system from pure observational quantities (the ‘dynamical clock’): in the same way as the engine of a chronometer advances a clock hand to measure the flow of time, in a similar way dynamical friction moves r_{min} within the cluster, measuring its dynamical age. Confirmation that this is indeed the case is provided by the tight correlations (see Fig. 4) obtained between the clock hand (r_{min}) and two theoretical estimates commonly used to measure the dynamical evolution timescales of a cluster, namely the central and the half-mass relaxation times, t_{rc} and t_{rh} , respectively¹ (see Supplementary Information), here expressed in units of the Hubble time (t_H). The best-fit relations to the data,

$$\log(t_{rc}/t_H) = -1.11 \log(r_{min}) - 0.78 \text{ (r.m.s.} = 0.32)$$

$$\log(t_{rh}/t_H) = -0.33 \log(r_{min}) - 0.25 \text{ (r.m.s.} = 0.23)$$

where r.m.s. is root mean square, can be assumed to be a preliminary calibration of the dynamical clock. Although t_{rc} and t_{rh} are indicative of the relaxation timescales at specific radial distances from the cluster centre, the dynamical clock here defined is much more sensitive to the global dynamical evolutionary stage reached by the system. In fact, the radial distribution of BSSs simultaneously probes all distances from the cluster centre, providing a measure of the overall dynamical evolution and a much finer ranking of dynamical ages. In the near future more realistic N -body simulations will provide a direct calibration of r_{min} as a function of the cluster’s dynamical age in billions of years.

Received 7 June; accepted 16 October 2012.

1. Meylan, G. & Heggie, D. C. Internal dynamics of globular clusters. *Annu. Rev. Astron. Astrophys.* **8**, 1–143 (1997).
2. Shara, M. M., Saffer, R. A. & Livio, M. The first direct measurement of the mass of a blue straggler in the core of a globular cluster: BSS19 in 47 Tucanae. *Astrophys. J.* **489**, L59–L63 (1997).

3. Hills, J. G. & Day, C. A. Stellar collisions in globular clusters. *Astrophys. J.* **17**, 87–93 (1976).
4. Sills, A., Adams, T., Davies, M. B. & Bate, M. R. High-resolution simulations of stellar collisions between equal-mass main-sequence stars in globular clusters. *Mon. Not. R. Astron. Soc.* **332**, 49–54 (2002).
5. McCrea, W. H. Extended main-sequence of some stellar clusters. *Mon. Not. R. Astron. Soc.* **128**, 147–155 (1964).
6. Sollima, A., Lanzoni, B., Beccari, G., Ferraro, F. R. & Fusi Pecci, F. The correlation between blue straggler and binary fractions in the core of Galactic globular clusters. *Astron. Astrophys.* **481**, 701–704 (2008).
7. Knigge, C., Leigh, N. & Sills, A. A binary origin for ‘blue stragglers’ in globular clusters. *Nature* **457**, 288–290 (2009).
8. Marin-Franch, A. et al. The ACS survey of Galactic globular clusters. VII. Relative ages. *Astrophys. J.* **694**, 1498–1516 (2009).
9. Dotter, A., Sarajedini, A. & Yang, S.-C. Globular clusters in the outer Galactic halo: AM-1 and Palomar 14. *Astron. J.* **136**, 1407–1414 (2008).
10. Ferraro, F. R. et al. Blue stragglers in the galactic globular clusters M3: evidence for two populations. *Astron. J.* **106**, 2324–2334 (1993).
11. Ferraro, F. R. et al. The pure noncollisional blue straggler population in the giant stellar system ω Centauri. *Astrophys. J.* **638**, 433–439 (2006).
12. Mapelli, M. et al. The contribution of primordial binaries to the blue straggler population in 47 Tucanae. *Astrophys. J.* **605**, L29–L32 (2004).
13. Mapelli, M. et al. The radial distribution of blue straggler stars and the nature of their progenitors. *Mon. Not. R. Astron. Soc.* **373**, 361–368 (2006).
14. Davies, M. B., Piotto, G. & de Angeli, F. Blue straggler production in globular clusters. *Mon. Not. R. Astron. Soc.* **349**, 129–134 (2004).
15. Geller, A. M. & Mathieu, R. D. A mass transfer origin for blue stragglers in NGC188 as revealed by half-solar-mass companions. *Nature* **478**, 356–359 (2011).
16. Binney, J. & Tremaine, S. *Galactic Dynamics* (Princeton Univ. Press, 1987).
17. Lanzoni, B. et al. The blue straggler population of the globular cluster M5. *Astrophys. J.* **663**, 267–276 (2007).
18. Mapelli, M. et al. Blue straggler stars in dwarf spheroidal galaxies. II. Sculptor and Fornax. *Mon. Not. R. Astron. Soc.* **396**, 1771–1782 (2009).
19. Monelli, M. et al. The ACS LCID Project. VII. The blue stragglers population in the isolated dSph galaxies Cetus and Tucana. *Astrophys. J.* **744**, 157–170 (2012).
20. Ferraro, F. R. et al. Two distinct sequences of blue straggler stars in the globular cluster. *Nature* **462**, 1028–1031 (2009).
21. Trager, S. C., Djorgovski, S. & King, I. R. in *Structure and Dynamics of Globular Clusters* (eds Djorgovski, S. G. & Meylan, G.) (Astronomical Society of the Pacific Conference Series Vol. 50) 347–355 (Astronomical Society of the Pacific, 1993).
22. Heggie, D. C. & Giersz, M. Monte Carlo simulations of star clusters. V. The globular cluster M4. *Mon. Not. R. Astron. Soc.* **389**, 1858–1870 (2008).
23. Ferraro, F. R. et al. The puzzling dynamical status of the core of the globular cluster NGC 6752. *Astrophys. J.* **595**, 179–186 (2003).
24. Lanzoni, B. et al. The surface density profile of NGC 6388: a good candidate for harboring an intermediate-mass black hole. *Astrophys. J.* **668**, L139–L142 (2007).
25. Renzini, A. & Buzzoni, A. in *Spectral Evolution of Galaxies* (eds Chiosi, C. & Renzini, A.) 195–231 (Reidel, 1986).
26. Bekki, K. & Freeman, K. C. Formation of ω Centauri from an ancient nucleated dwarf galaxy in the young Galactic disc. *Mon. Not. R. Astron. Soc.* **346**, L11–L15 (2003).
27. Djorgovski, S. in *Structure and Dynamics of Globular Clusters* (eds Djorgovski, S. G. & Meylan, G.) (Astronomical Society of the Pacific Conference Series Vol. 50) 373–382 (Astronomical Society of the Pacific, 1993).
28. King, I. R. The structure of star clusters. III. Some simple dynamical models. *Astron. J.* **71**, 64–75 (1966).
29. Ferraro, F. R. et al. The giant, horizontal, and asymptotic branches of Galactic globular clusters. I. The catalog, photometric observables, and features. *Astron. J.* **118**, 1738–1758 (1999).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors dedicate this paper to the memory of co-author Bob Rood, a pioneer in the theory of the evolution of low mass stars and a friend who shared our enthusiasm for the BSS topic, who passed away on 2 November 2011. This research is part of the project COSMIC-LAB funded by the European Research Council (under contract ERC-2010-AdG-267675). G.B. acknowledges the European Community’s Seventh Framework Programme under grant agreement no. 229517. F.R.F. acknowledges support from the ESO Visiting Scientist Programme. This research is based on data acquired with the NASA/ESA HST, under programmes GO-11975, GO-10524, GO-8709, GO-6607 and GO-5903 at the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. The research is also based on data collected at the ESO telescopes under programmes 62.L-0354, 64.L-0439, 59.A-002(A), 69.D-0582(A), 079.D-0220(A) and 079.D-0782(A), and made use of the ESO/ST-ECF Science Archive facility, which is a joint collaboration of the European Southern Observatory and the Space Telescope – European Coordinating Facility.

Author Contributions F.R.F. designed the study and coordinated the activity. E.D., G.B., R.C., B.L., N.S. and A.M. analysed the data. M.P. and P.M. developed N -body simulations. F.R.F. and B.L. wrote the paper. E.V., A.S., S.S., M.M. and R.T.R. critically contributed to discussion and presentation of paper. All authors contributed to discussion of the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.R.F. (francesco.ferraro3@unibo.it).

Evaporative cooling of the dipolar hydroxyl radical

Benjamin K. Stuhl¹, Matthew T. Hummon¹, Mark Yeo¹, Goulven Quémener^{1†}, John L. Bohn¹ & Jun Ye¹

Atomic physics was revolutionized by the development of forced evaporative cooling, which led directly to the observation of Bose–Einstein condensation^{1,2}, quantum-degenerate Fermi gases³ and ultracold optical lattice simulations of condensed-matter phenomena⁴. More recently, substantial progress has been made in the production of cold molecular gases⁵. Their permanent electric dipole moment is expected to generate systems with varied and controllable phases^{6–8}, dynamics^{9–11} and chemistry^{12–14}. However, although advances have been made¹⁵ in both direct cooling and cold-association techniques, evaporative cooling has not been achieved so far. This is due to unfavourable ratios of elastic to inelastic scattering¹³ and impractically slow thermalization rates in the available trapped species. Here we report the observation of microwave-forced evaporative cooling of neutral hydroxyl (OH^{*}) molecules loaded from a Stark-decelerated beam into an extremely high-gradient magnetic quadrupole trap. We demonstrate cooling by at least one order of magnitude in temperature, and a corresponding increase in phase-space density by three orders of magnitude, limited only by the low-temperature sensitivity of our spectroscopic thermometry technique. With evaporative cooling and a sufficiently large initial population, much colder temperatures are possible; even a quantum-degenerate gas of this dipolar radical (or anything else it can sympathetically cool) may be within reach.

Evaporative cooling of a thermal distribution¹⁶ is, in principle, very simple: by selectively removing particles with energies much greater than the average total energy per particle, the temperature decreases. In the presence of elastic collisions, the high-energy tail is repopulated and so may repeatedly be selectively trimmed, allowing the removal of a great deal of energy at low cost in particle number. This process may be started as soon as the thermalization rate is fast enough to be practical and continued until its cooling power is balanced by the heating rate from inelastic collisions. It generally yields temperatures deep into the quantum-degenerate regime (far below the recoil limit of optical cooling), and in a few cases has reached quantum degeneracy without requiring optical pre-cooling^{17,18}.

The key metric for evaporation is therefore the ratio of two time-scales. The first is the rate of elastic collisions, which rethermalize the distribution, while the second is the rate at which particles are lost from the trap for reasons other than their being deliberately removed—for example, the rates of inelastic scattering and background gas collisions. For a long time, both theoretical^{14,19,20} and experimental^{13,21,22} work seemed to show a generically poor value of this ratio across multiple molecular systems; this has led to a general belief that evaporative cooling is unfavourable for molecules¹⁵. As no trapped molecular system has achieved sufficiently rapid thermalization, there has been a lack of experiments to test this expectation—although some recent theoretical results have challenged it^{23–25}.

Hydroxyl (OH^{*}, referred to here as OH) would not, at first glance, seem to be a promising candidate for evaporative cooling. Its open-shell ²Π_{3/2} ground state and its propensity towards hydrogen bonding create a large anisotropy in the OH–OH interaction potential, which would intuitively produce a large inelastic scattering rate. Chemical

reactions are also possible, via the OH + OH → H₂O + O pathway; it is unclear whether this reaction has an activation energy barrier²⁶. It is thus perhaps surprising that the elastic collision rate actually exceeds the inelastic rate, allowing evaporative cooling. However, our experimental observation is unambiguous, and is further supported by quantum scattering calculations based on the long-range dipole–dipole interaction between the molecules^{23,27} considering all of the fine-structure states of the rotational ground state. This analysis considers only elastic collision or inelastic relaxation to lower energy states, as the long-range interactions appear to fully dominate over short-range effects, such as chemical reactions.

In its ground state, OH has a ²Π_{3/2} electronic character, with the lowest rotational level having a total angular momentum, excluding nuclear spin, of $J = 3/2$. The electronic orbital angular momentum couples to the rotational angular momentum to split the two opposite-parity states within $J = 3/2$ by a Λ -doubling of ~ 1.667 GHz; the upper parity state is labelled $|f\rangle$ and the lower $|e\rangle$. Hydroxyl is both paramagnetic, with a molecule-fixed moment of $2\mu_B$ (μ_B is the Bohr magneton), and electrically polar, with a dipole moment of 1.67 Debye (5.57×10^{-30} C m). The Zeeman spectrum of OH is shown in Fig. 1a. Our magnetic trap²⁸ is loaded with molecules in the uppermost $|f; M_J = +3/2\rangle$ state, where M_J is the laboratory projection of J .

The results of our scattering calculations for $|f; +3/2\rangle$ molecules in zero electric field are shown in Fig. 1b. The elastic cross-section dominates the inelastic one for low energies: at a collision energy of $E_c = 50$ mK, the ratio of elastic over inelastic is $R = 5$ in a 50 mT magnetic field (B) and $R = 18$ in 150 mT, whereas at lower energy $E_c = 5$ mK the ratio increases to $R = 23$ and 137, respectively. Because the collisions occur in a quadrupole magnetic trap where B is inhomogeneous, the inset of Fig. 1b shows the cross-sections as a function of B at $E_c = 50$ mK. This demonstrates that inelastic processes are even further suppressed at $B > 50$ mT, in agreement with previous analysis²³.

These scattering results can be interpreted by the emergence of an effective repulsive van der Waals interaction between the two molecules. In zero electric field, the effective interaction V_{vdW} at distance r in a scattering channel m can be evaluated in second-order perturbation theory by

$$V_{\text{vdW}}(r) = \sum_n \frac{|\langle m | V_{\text{dd}}(r) | n \rangle|^2}{E_m - E_n} \equiv \frac{C_6}{r^6} \quad (1)$$

where V_{dd} is the electric dipole–dipole interaction between different scattering channels m and n , and is non-vanishing only between molecular states of distinct parity; and E_m and E_n are the asymptotic energies of the relevant scattering channels. For the initial molecule–molecule channel of interest $|f; +3/2\rangle |f; +3/2\rangle$, the other fine-structure channels are lower in energy ($E_n < E_m$) and repel this energy upward. Thus, the C_6 coefficient is positive, as illustrated in Fig. 1c (highlighted as red lines in the inset) for $B = 50$ mT, for the highest energy channel. The contribution from the next rotational state $J = 5/2$ of the ²Π_{3/2} manifold is too high in energy ($E_m - E_{J=5/2} \approx -100$ K) to give an appreciable attractive contribution at long-range. Hence, for low collision energies, the scattering of OH molecules is dominated by

¹JILA, National Institute of Standards and Technology and University of Colorado, Department of Physics, University of Colorado, Boulder, Colorado 80309, USA. [†]Present address: Laboratoire Aimé Cotton, Université Paris-Sud, CNRS, Bâtiment 505, 91405 Orsay, France.

the long-range interaction rather than the short-range structure of the potential surface.

Given an effective, repulsive C_6 coefficient in the incident channel, the elastic cross-section σ^{el} can be approximated semi-classically at energies above the threshold regime²⁹ as

$$\sigma^{\text{el}} = \frac{\pi^{11/5} (\Gamma(5/2)/\Gamma(3))^{2/5}}{\sin(\pi/5) \Gamma(2/5)} \left(\frac{\bar{C}_6}{\hbar} \right)^{2/5} \left(\frac{2E_c}{m_{\text{red}}} \right)^{-1/5} \quad (2)$$

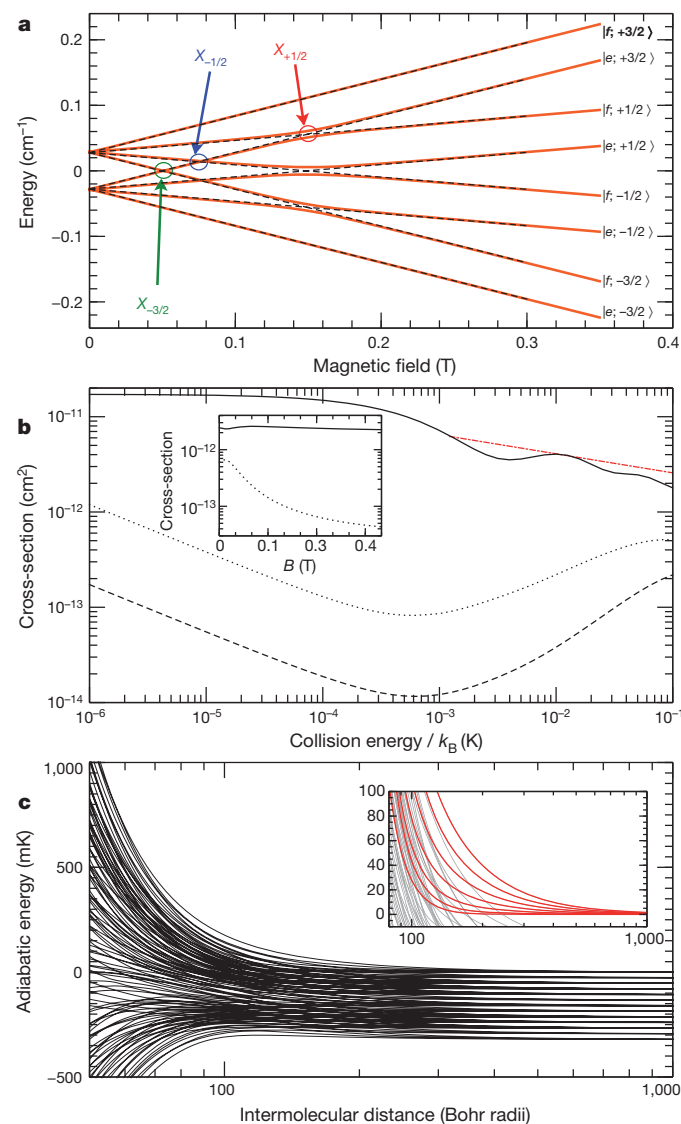


Figure 1 | Ground-state structure and scattering theory of OH. **a**, The ground-state Λ -doublet and Zeeman structure of OH. Dashed black lines are energy levels in the absence of any electric field. In the presence of an electric field, the level crossings become avoided (solid orange lines demonstrate a 500 V cm^{-1} field orthogonal to the magnetic field). X_i label the crossings of the $|e; +3/2\rangle$ state with the $|f; M_J = i\rangle$ states; these crossings allow $|e; +3/2\rangle$ molecules to escape the trap in the presence of an electric field. Molecules are loaded into the magnetic trap in the $|f; +3/2\rangle$ state (bold). **b**, Cross-sections in zero electric field as a function of collision energy: elastic cross-section in a 50 mT magnetic field (solid line), inelastic in 50 mT (dotted), and inelastic in 150 mT (dashed). The red dash-dotted line is a semi-classical expression given by equation (2). Inset, elastic (solid line) and inelastic (dotted) cross-sections as a function of magnetic field at $E_c = 50 \text{ mK}$. **c**, Adiabatic energies as a function of the intermolecular distance at $B = 50 \text{ mT}$ for every combination of molecular states using $\ell = [0-12]$ even partial waves. Inset, magnified views of the repulsive van der Waals interaction for the experimentally relevant case of two colliding $|f; +3/2\rangle$ molecules (red lines); quantities plotted on the axes are the same as in the main panel.

where $m_{\text{red}} \approx 8.5 \text{ AMU}$ is the reduced mass of two OH molecules and $\bar{C}_6 \approx 9 \times 10^4 \text{ a.u.}$ is the calculated isotropic C_6 . Γ is the Gamma function. (We contrast this \bar{C}_6 with values of $(2-8) \times 10^3 \text{ a.u.}$ for the alkali metal atoms³⁰.) Equation (2), plotted as a red dash-dotted line on Fig. 1b, only slightly overestimates the numerical results for the elastic cross-section but gives a proper magnitude and trend in $E_c^{-1/5}$.

We observed inelastic scattering in the presence of a large electric field (Fig. 2b), demonstrating the presence of two-body collisions in our trapped sample. Motivated by the prediction of a favourable R , we then undertook an experiment to implement microwave-forced evaporative cooling. (A detailed description of the experiment is provided in Methods.) The $|f; +3/2\rangle \rightarrow |e; +3/2\rangle$ microwave transition has a small but non-zero differential Zeeman shift³¹, redshifting by 26.6 kHz mT^{-1} . In the presence of a small electric field, $|e; +3/2\rangle$ molecules can escape the trap through the avoided crossings labelled X_i in Fig. 1a³², while inelastic losses of $|f\rangle$ molecules remain unmeasurably slow. A brief microwave pulse to selectively transfer $|f; +3/2\rangle$ -state molecules to $|e; +3/2\rangle$, followed by a longer period with an electric field present to eject $|e; +3/2\rangle$ molecules from the trap, is therefore a field- (position-) selective method to remove

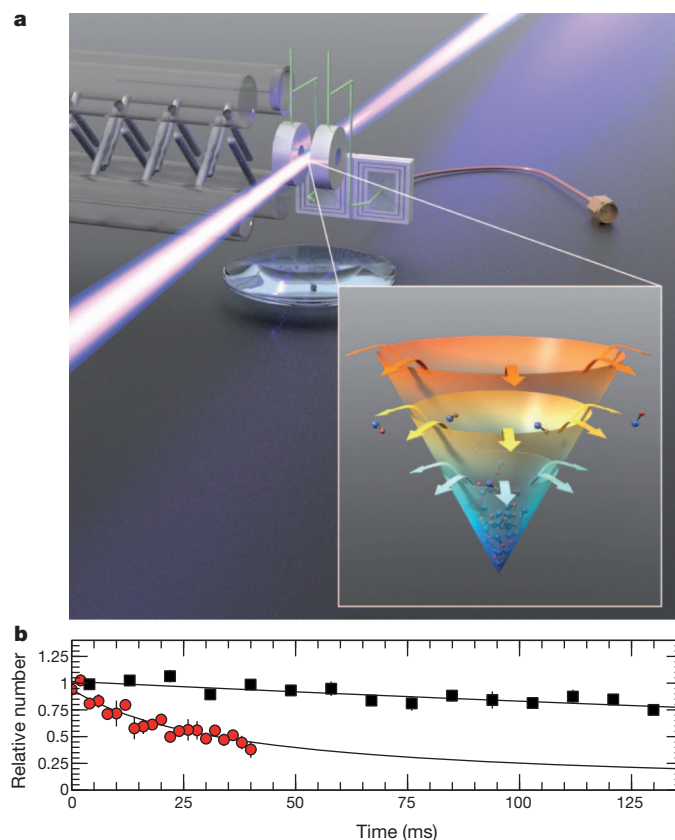


Figure 2 | Trap system and inelastic collisions. **a**, Schematic of the Stark decelerator³³ (left) and magnetic trap (centre rings) system. The d.c.-blocking capacitor (rear) decouples the high voltage used for trap loading from the coaxial cable carrying the microwaves used for spectroscopy and evaporation, while the laser and the lens below the magnets provide fluorescence detection of the trapped molecules. Inset, an artist's impression of evaporative cooling. As the trap depth is successively lowered by the r.f. knife, the hottest molecules escape (coloured arrows) and the remainder rethermalize to a colder temperature. Temperature is symbolized by colour, with red denoting the highest temperature regions of the trap and blue the lowest. **b**, Relative molecule number versus trap hold time, demonstrating electric-field-induced collisions at 45 mK . Black squares, consistent with pure one-body loss, are with no applied electric field; red circles are well-fitted by pure two-body inelastic loss induced by a large applied electric field of $3,040 \text{ V cm}^{-1}$. (The field used in the r.f. knife is only 240 V cm^{-1} , for comparison.) Error bars, one standard error.

molecules from the magnetic trap (see Methods). This yields what is commonly called a radio frequency (r.f.) knife (Fig. 2a). A Zeeman depletion spectrum can be acquired by using the knife to remove molecules at a set of fixed frequencies and measuring the fractional depletion ξ ; this yields a histogram of relative molecule number versus B . We fitted this spectrum with a modified thermal distribution

$$\xi(B)dB = \xi_0 B^2 dB \times \exp\left[-\frac{\mu B}{k_B T}\right] \times \begin{cases} 1, & \text{if } B > 49.6 \text{ mT} \\ \exp\left\{-\frac{\mu[(49.6 \text{ mT}) - B]}{k_B T}\right\}, & \text{otherwise} \end{cases} \quad (3)$$

where ξ_0 is a fitting coefficient, $B^2 dB \propto r^2 dr$ is the volume element for a three-dimensional quadrupole trap, $\mu = 1.2 \mu_B$ is the magnetic moment of the $|f; +3/2\rangle$ state, B is the magnetic field strength implied by the microwave frequency, k_B is Boltzmann's constant, and T is the fitted temperature. The first term in the product is the simple Boltzmann expression for the molecule number as a function of B , while the second is a correction for the fact that $|e; +3/2\rangle$ molecules only disappear when they go through one of the X_i crossings. Specifically, while molecules transferred at fields above 49.6 mT (the known location of $X_{-3/2}$) are always energetically able to reach one of the X_i crossings and thus disappear, of the molecules transferred at lower fields only those with enough kinetic energy to fly up the trap potential to $X_{-3/2}$ can escape the trap. This implies an additional Boltzmann factor $\exp[-\mu\Delta B/k_B T]$ in the probability of those molecules' disappearance.

With a B -selective technique for removing molecules, we easily implement forced evaporation by moving the edge of the r.f. knife from some large initial value of B towards zero, at a rate slow enough that the distribution remains in thermal equilibrium. Six different Zeeman spectra are shown in Fig. 3, demonstrating both forced evaporation and forced anti-evaporation, where for the latter the knife is held fixed at some low B_{knife} ($\mu B_{\text{knife}} < k_B T$). The temperature of the molecules in the trap immediately after loading is 51 mK; left unperturbed, the sample free-evaporates down to 45 mK. We have forced evaporative cooling by an order of magnitude down to 5.1 mK, while forced anti-evaporation can overcome the free evaporation and increase the temperature to 54 mK. The limit of 5.1 mK is attained approximately when the r.f. knife edge reaches $X_{-3/2}$; although further reductions in temperature are possible, the exponential suppression of the spectroscopic signal below $X_{-3/2}$ renders our current system unable to measure temperatures lower than this. The plots of Fig. 3 are all normalized so that the area under the spectroscopic curve is proportional to the total $|f\rangle$ -state fluorescence signal. Thus, the increase in signal height at low B in Fig. 3c–f is direct evidence of increasing phase-space density (PSD).

The apparent negative signal in Fig. 3e, f can be fully fitted by assuming the presence of accumulated, trapped $|e\rangle$ -state molecules in thermal equilibrium with the visible $|f\rangle$ -state ones. Because $|e; +3/2\rangle$ molecules are totally trapped if they do not have enough kinetic energy to reach $X_{-3/2}$, they will appear, rather than disappear, during the microwave spectroscopy and contribute to the total depletion signal with a negative sign. Fitting the curves in this fashion gives even colder temperatures of 6.8 and 3.5 mK for Fig. 3e and f, respectively. As the appearance of low-energy, trapped $|e\rangle$ -state molecules would also constitute direct evidence of evaporation, we undertook a direct search and indeed detected them in laser-induced fluorescence.

The high speed of the observed evaporation should be emphasized. In comparison to typical cooling rates of $d(\log T)/dt \approx 0.5 \text{ s}^{-1}$, we cool by an order of magnitude in only 70 ms, for a rate of 33 s^{-1} . This implies elastic collision rates of the order of $100\text{--}1,000 \text{ s}^{-1}$, comparable to our trap frequency³² of $\sim 1,400 \text{ s}^{-1}$. As we are able to set a bound on the inelastic loss rate of $<2 \text{ s}^{-1}$ (as shown in Fig. 2b), this implies a

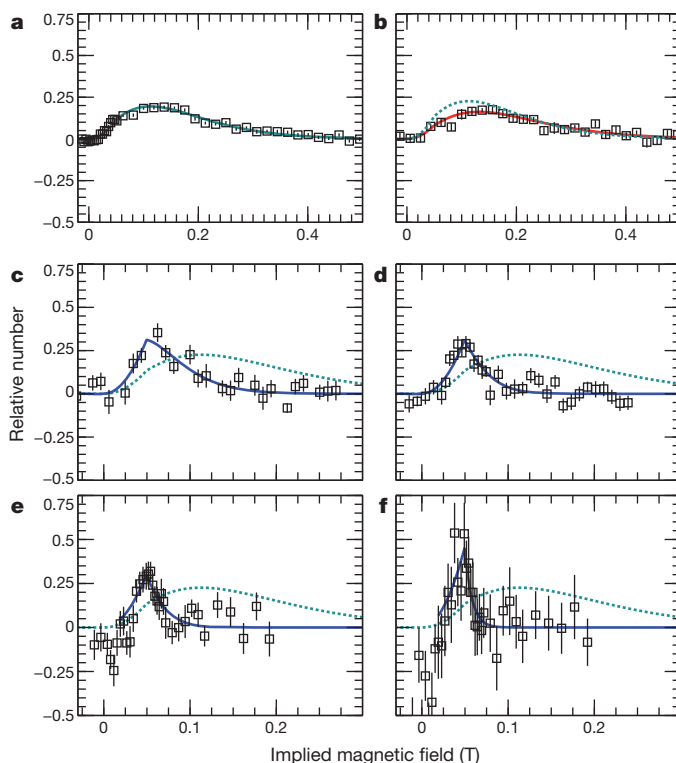


Figure 3 | Microwave spectra illustrating six different final temperatures. Black squares are data (error bars, one standard error), and solid lines are fits to the sensitivity-corrected Boltzmann form of equation (3): **a**, the unperturbed 45 mK distribution; **b**, anti-evaporation to 54 mK; **c–f**, forward evaporation to 18 mK (**c**), 12 mK (**d**), 9.8 mK (**e**) and 5.1 mK (**f**). Dotted lines reproduce the fit from **a**, shown for comparison. Note that the x -axis scale differs between **a–b** and **c–f**; all curves have been shifted vertically to have a zero baseline.

distribution-averaged $R \gtrsim 50\text{--}500$, consistent with the B -field dependence of R shown in Fig. 1b.

We quantify the behaviour of the evaporation by a set of power-law scaling relations¹⁶, three of which are plotted in Fig. 4. The average energy removed per molecule, $\eta \equiv \mu B_{\text{knife}}/k_B T$, is the first metric of interest: we observe a ratio $\eta = 5.6$, as shown in Fig. 4a. (For scale, $\eta = 5\text{--}10$ is considered reasonable in atomic evaporation¹⁶.) Such a relatively low value of η is unsurprising, given both the possible low ($<1,000$) value of the elastic-to-inelastic ratio R and the fact that molecules are only actually lost when they cross the spatial regions corresponding to the X_i crossings. This reduces the selectivity of the r.f. knife, as molecules transferred to the $|e\rangle$ -state may recollide and rethermalize before they find their way to a crossing. The other metrics are the behaviour of temperature and relative PSD as functions of remaining molecule number, shown in Fig. 4b. The efficiency of evaporation is determined by the number of molecules remaining at a given temperature and PSD, that is by $\alpha \equiv d(\log T)/d(\log N)$ and $\gamma \equiv -d(\log \text{PSD})/d(\log N)$. We measure $\alpha = 1.26$ and so, using the fact that density scales as T^{-3} for our three-dimensional quadrupole trap, we find $\gamma = 4.7$. Because R is worse at 15 mK than any other temperature down to 1 μK , linear extrapolation suggests that it would take roughly a factor of 200 reduction in number to increase PSD by 10^{10} , while the repulsive s -wave barrier would serve to stabilize the cloud as it gets close to the quantum regime.

We calculate both α and γ assuming that our fluorescence signal is linear with total molecule number, justified by the optical power broadening being larger than the Zeeman broadening of the trap. It is very difficult to determine the sensitivity of pulsed-laser-induced fluorescence *in situ*, so we use the observed collision rate to estimate a lower bound on the density in our trap: assuming a unitarity-limited scattering rate β_u with a maximum collisional angular momentum of

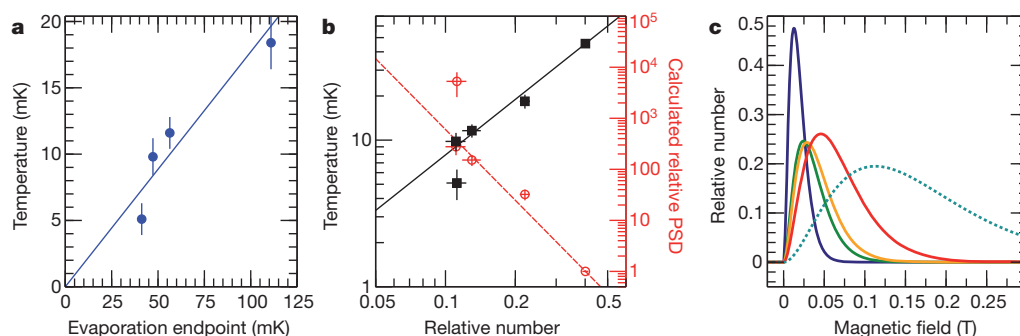


Figure 4 | Scaling relations observed in the evaporative cooling of OH.

a, Final observed temperature versus effective trap depth at evaporation ramp end. **b**, Final temperature (solid black, left axis) and calculated relative phase-space density (PSD; dashed red, right axis) versus remaining relative molecule number. Error bars, one standard error; phase space density is calculated

$12\hbar$ (and the elastic scattering rate in Fig. 1b is only a factor of three below this value), an elastic collision rate of $\beta_u n_0 = 300 \text{ s}^{-1}$ implies a peak density $n_0 \approx 5 \times 10^{10} \text{ cm}^{-3}$. (Integration of a Boltzmann distribution with this peak density over the trap suggests a total number of $\sim 10^6$ molecules in the free-evaporated trap sample and an initial peak PSD of 3×10^{-10} .) This density is high enough to permit the use of absorption imaging techniques to directly visualize the trap distribution in the future. Because imaging allows direct, non-spectroscopic measurement of both density and temperature, it will enable us to quantify further reductions in temperature. With the addition of a mechanism to prevent Majorana loss¹⁶, the favourability of R down to microkelvin temperatures indicates that even Bose–Einstein condensation of OH may be feasible.

METHODS SUMMARY

Our Stark decelerator and permanent magnet trap have been described elsewhere^{28,33} and are illustrated in Fig. 2a. See also the Supplementary Information. Briefly, OH molecules are formed by an electric discharge through a saturated mixture of water vapour in 150 kPa of krypton, supersonically expanded through a pulsed valve. The gas packet is skimmed, focused by an electrostatic hexapole, and Stark-decelerated to 34 m s^{-1} . The slowed packet is then stopped at the centre of the magnetic quadrupole trap by a high-voltage field applied between the permanent magnets, and thus loaded into the magnetic trap.

Once trapped, the molecules are allowed to settle briefly (of the order of 5 ms) before evaporation begins. The evaporation sequence consists of alternating microwave ($80 \mu\text{s}$ with 0 dBm at the vacuum feed-through) and d.c. bias field ($\sim 650 \mu\text{s}$ at 240 V cm^{-1}) pulses: the microwave pulse selectively transfers hot molecules from the $|f\rangle$ -state to the $|e\rangle$ -state, while the d.c. bias destabilizes the $|e\rangle$ -state so that those molecules are lost from the trap. The microwave frequency is ramped along an exponential curve towards zero trap depth, truncated at the desired final depth.

After evaporation, the spectroscopy sequence is executed. It is similar to the evaporation, except that rather than slowly ramping a set of quasi-single-frequency pulses, each pulse is rapidly swept through the same narrow (50–500 kHz) frequency band with an additional sine-wave amplitude modulation ($+10 \text{ dBm}$ peak power at the feed-through) so as to induce adiabatic rapid passage (ARP) transference³² of all the $|f\rangle$ -state molecules within the frequency band to the $|e\rangle$ -state. A d.c. bias field again rejects the $|e\rangle$ -state molecules from the trap. The number of ARP pulses applied is generally between 5 and 30, over a 2–10 ms spectroscopy sequence. The final molecule number is then detected by pulsed laser-induced fluorescence using the $282 \text{ nm } X^2\Pi_{3/2}, v''=0 \rightarrow A^2\Sigma, v'=1$ transition.

Full Methods and any associated references are available in the online version of the paper.

Received 11 September; accepted 26 October 2012.

- Anderson, M. H., Ensher, J. R., Matthews, M. R., Wieman, C. E. & Cornell, E. A. Observation of Bose–Einstein condensation in a dilute atomic vapor. *Science* **269**, 198–201 (1995).
- Davis, K. B. *et al.* Bose–Einstein condensation in a gas of sodium atoms. *Phys. Rev. Lett.* **75**, 3969–3973 (1995).

assuming a three-dimensional quadrupole trap geometry. **c**, Molecular Boltzmann distributions implied by the fitted curves of Fig. 3a (dashed) and c–f (right to left). Distributions are normalized so that the area under each curve is proportional to the total $|f\rangle$ -state fluorescence observed at that temperature.

- DeMarco, B. & Jin, D. S. Onset of Fermi degeneracy in a trapped atomic gas. *Science* **285**, 1703–1706 (1999).
- Bakr, W. S. *et al.* Probing the superfluid-to-Mott insulator transition at the single-atom level. *Science* **329**, 547–550 (2010).
- Ni, K.-K. *et al.* A high phase-space-density gas of polar molecules. *Science* **322**, 231–235 (2008).
- Pupillo, G. *et al.* Cold atoms and molecules in self-assembled dipolar lattices. *Phys. Rev. Lett.* **100**, 050402 (2008).
- Baranov, M. A., Micheli, A., Ronen, S. & Zoller, P. Bilayer superfluidity of fermionic polar molecules: many-body effects. *Phys. Rev. A* **83**, 043602 (2011).
- Levinson, J., Cooper, N. R. & Shlyapnikov, G. V. Topological $p_x + ip_y$ superfluid phase of fermionic polar molecules. *Phys. Rev. A* **84**, 013603 (2011).
- Barnett, R., Petrov, D., Lukin, M. & Demler, E. Quantum magnetism with multicomponent dipolar molecules in an optical lattice. *Phys. Rev. Lett.* **96**, 190401 (2006).
- Büchler, H. P. *et al.* Strongly correlated 2D quantum phases with cold polar molecules: controlling the shape of the interaction potential. *Phys. Rev. Lett.* **98**, 060404 (2007).
- Gorshkov, A. V. *et al.* Tunable superfluidity and quantum magnetism with ultracold polar molecules. *Phys. Rev. Lett.* **107**, 115301 (2011).
- Ospelkaus, S. *et al.* Quantum-state controlled chemical reactions of ultracold potassium-rubidium molecules. *Science* **327**, 853–857 (2010).
- Ni, K.-K. *et al.* Dipolar collisions of polar molecules in the quantum regime. *Nature* **464**, 1324–1328 (2010).
- Quémener, G. & Julienne, P. S. Ultracold molecules under control! *Chem. Rev.* **112**, 4949–5011 (2012).
- Carr, L. D., DeMille, D., Krems, R. V. & Ye, J. Cold and ultracold molecules: science, technology and applications. *N. J. Phys.* **11**, 055049 (2009).
- Ketterle, W. & VanDruten, N. Evaporative cooling of trapped atoms. *Adv. At. Mol. Opt. Phys.* **37**, 181–236 (1996).
- Fried, D. G. *et al.* Bose–Einstein condensation of atomic hydrogen. *Phys. Rev. Lett.* **81**, 3811–3814 (1998).
- Doret, S. C., Connolly, C. B., Ketterle, W. & Doyle, J. M. Buffer-gas cooled Bose–Einstein condensate. *Phys. Rev. Lett.* **103**, 103005 (2009).
- Lara, M., Bohn, J. L., Potter, D. E., Soldan, P. & Hutson, J. M. Ultracold Rb–OH collisions and prospects for sympathetic cooling. *Phys. Rev. Lett.* **97**, 183201 (2006).
- Zuchowski, P. S. & Hutson, J. M. Low-energy collisions of NH_3 and ND_3 with ultracold Rb atoms. *Phys. Rev. A* **79**, 062708 (2009).
- Campbell, W. *et al.* Mechanism of collisional spin relaxation in $^3\Sigma$ molecules. *Phys. Rev. Lett.* **102**, 013003 (2009).
- Parazzoli, L. P., Fitch, N. J., Zuchowski, P. S., Hutson, J. M. & Lewandowski, H. J. Large effects of electric fields on atom-molecule collisions at millikelvin temperatures. *Phys. Rev. Lett.* **106**, 193201 (2011).
- Ticknor, C. & Bohn, J. L. Influence of magnetic fields on cold collisions of polar molecules. *Phys. Rev. A* **71**, 022709 (2005).
- Janssen, L. M. C., Zuchowski, P. S., van der Avoird, A., Groenenboom, G. C. & Hutson, J. M. Cold and ultracold NH–NH collisions in magnetic fields. *Phys. Rev. A* **83**, 022713 (2011).
- Suleimanov, Y. V., Tscherebul, T. V. & Krems, R. V. Efficient method for quantum calculations of molecule-molecule scattering properties in a magnetic field. *J. Chem. Phys.* **137**, 024103 (2012).
- Manion, J. A. *et al.* NIST Chemical Kinetics Database. <http://kinetics.nist.gov/> (NIST Standard Reference Database 17, Web version 70, Release 1.43, Data version 2008.12).
- Avdeenkov, A. V. & Bohn, J. L. Collisional dynamics of ultracold OH molecules in an electrostatic field. *Phys. Rev. A* **66**, 052718 (2002).
- Sawyer, B. C., Stuhl, B. K., Wang, D., Yeo, M. & Ye, J. Molecular beam collisions with a magnetically trapped target. *Phys. Rev. Lett.* **101**, 203203 (2008).
- Child, M. S. *Molecular Collision Theory* 76 (Dover Publications, 1996).
- Derevianko, A., Johnson, W. R., Safronova, M. S. & Babb, J. F. High-precision calculations of dispersion coefficients, static dipole polarizabilities, and atom-wall interaction constants for alkali-metal atoms. *Phys. Rev. Lett.* **82**, 3589–3592 (1999).

31. Lev, B. L. *et al.* OH hyperfine ground state: from precision measurement to molecular qubits. *Phys. Rev. A* **74**, 061402 (2006).
32. Stuhl, B. K., Yeo, M., Sawyer, B. C., Hummon, M. T. & Ye, J. Microwave state transfer and adiabatic dynamics of magnetically trapped polar molecules. *Phys. Rev. A* **85**, 033427 (2012).
33. Bochinski, J. R., Hudson, E. R., Lewandowski, H. J., Meijer, G. & Ye, J. Phase space manipulation of cold free radical OH molecules. *Phys. Rev. Lett.* **91**, 243001 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Cornell for discussions and B. Baxley for artistic contributions. We acknowledge funding from the NSF Physics Frontier Center, DOE, AFOSR (MURI), DARPA and NIST.

Author Contributions B.K.S., M.T.H., M.Y. and J.Y. designed and participated in the experiment, and discussed and interpreted the results. B.K.S. ran the day-to-day experiment and collected all the data. G.Q. and J.L.B. constructed the theory. B.K.S. and J.Y. first outlined the manuscript, and B.K.S. and G.Q. wrote the first draft. All authors discussed the results and contributed to the preparation of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.Y. (ye@jila.colorado.edu).

METHODS

Stark deceleration and magnetic trapping. Our Stark decelerator and permanent magnet trap have been described elsewhere^{28,33} and are illustrated both in Fig. 2a and with more technical detail in Supplementary Information. The molecular beam is formed by a supersonic expansion of a saturated mixture of water vapour in 150 kPa of krypton through a pulsed valve; a fraction of the water is cracked by a 1,250 V electric discharge (seeded by a hot-cathode ionization gauge) to produce OH. The gas packet is skimmed, focused by an electrostatic hexapole, and Stark-decelerated to 34 m s^{-1} . The slowed packet is then stopped at the centre of the magnetic quadrupole trap by a $\sim 61 \text{ kV cm}^{-1}$ electric field applied between the permanent magnets; turning off that field leaves the molecules with zero centre-of-mass velocity and confined by the magnetic trap. The magnetic trap field is created by a pair of NdFeB permanent magnets (grade N42SH), each of which is a (4 mm ID) \times (12 mm OD) \times (4 mm thick) ring; this yields a maximum gradient of $\sim 2 \text{ T cm}^{-1}$ in the strongly confining direction. The high voltage for loading the trap, the low voltage d.c. bias for ejecting $|e\rangle$ -state molecules, and the microwave fields are all directly applied to the surface plating of the permanent magnets.

Detection. We observe the presence of OH molecules through pulsed-laser induced fluorescence (LIF). The molecules are excited on the $A^2\Sigma_{1/2}(v=1) \leftarrow X^2\Pi_{3/2}(v=0)$ transition at 282 nm, and subsequently decay with $\sim 70\%$ efficiency by $A^2\Sigma_{1/2}(v=1) \rightarrow X^2\Pi_{3/2}(v=1)$ around 313 nm. This large Stokes shift enables strong rejection of the excitation light by a stack of optical filters atop the photomultiplier tube (PMT) detector. The fluorescence is collected in vacuum by a single collimating lens located one focal length away from the trap centre, and is subsequently focused onto the PMT. Unfortunately, neither the fluorescence collection efficiency nor the laser excitation efficiency are easily calculable, so LIF yields only a relative number measurement rather than an absolute one. LIF is parity-state selective, since the upper $A^2\Sigma_{1/2}$ state does not have a parity doublet: the $|e\rangle$ and $|f\rangle$ ground states must couple to different, spectroscopically resolvable excited rotational levels.

Spectroscopic thermometry. In order to determine the temperature of the trapped distribution, we use a microwave depletion spectroscopy technique. A series of adiabatic rapid passage (ARP)³² pulses are repeated over a single narrow (50–500 kHz) band to transfer Zeeman-selected molecules from $|f; +3/2\rangle$ to $|e; +3/2\rangle$, alternated with longer pulses of a d.c. electric field (390 V cm^{-1}) to eject the transferred $|e\rangle$ -state molecules from the trap through the X_i avoided crossings of Fig. 1a. The ARP pulse duration is generally $40 \mu\text{s}$, while the d.c. bias pulses vary from 200 to $1,000 \mu\text{s}$ over a typical 10 ms, 17-repetition sequence. Multiple ARP pulses are used to enhance the contrast of the depletion, as a single

ARP with +10 dBm peak power at the vacuum feed-through (limited by the maximum acceptable power-broadening of the microwave transition) yields less than 10% peak depletion. Since the ARPs are all at a single, fixed centre frequency, there is virtually no net evaporation or anti-evaporation from the repetitions; the only effect is to enhance the contrast.

In order to accumulate a full spectrum, many repetitions of the experiment are required: each measurement involves a new packet, which is slowed, trapped and evaporated, then spectroscopically depleted and finally detected by LIF, which destroys the sample. Each data point in the plots of Fig. 3 is an average of many repetitions of the experiment, normalized by independent undepleted population measurements. Specifically, the relative depletion D_{rel} is given by

$$D_{\text{rel}} = 1 - \frac{\text{molecules with depletion ON}}{\text{molecules with depletion OFF}}$$

Evaporation. Once trapped, the molecules are allowed to settle briefly (of the order of 5 ms) before evaporation begins. The evaporation sequence consists of alternating microwave ($80 \mu\text{s}$ with 0 dBm at the vacuum feed-through) and d.c. electric field ($\sim 650 \mu\text{s}$ at 240 V cm^{-1}) pulses, though in this case with no substantial frequency chirp within any single pulse. The microwave frequency is gradually ramped along an exponential curve towards zero trap depth, truncated at the desired final depth. The entire evaporation sequence was completed in 70 ms for all of Fig. 3c–f, so the effective ramp rate varies between the different final temperatures.

It is common in atomic experiments to use a magnetic-dipole transition to directly couple magnetically trapped states to untrapped ones, but that is not an option at the initial high temperatures in our trap. At 50 mK, hot molecules sample magnetic fields of over 300 mT, implying that the differential Zeeman shift between the $|f; +3/2\rangle$ and the $|f; +1/2\rangle$ states is over 3.3 GHz as can be seen in Fig. 1a. Since this is much larger than the Λ -doublet splitting, attempts to use drive this transition will also drive $|f\rangle \rightarrow |e\rangle$ transitions at other locations in the trap, removing the specificity of the r.f. knife. The $|f; +3/2\rangle \rightarrow |e; 3/2\rangle$ electric-dipole transition, by contrast, can be driven with low enough power as to make parasitic magnetic-dipole transitions a non-issue. Once the sample is cold enough that the r.f. knife has passed through $X_{+3/2}$ and the magnetic-dipole transition is far red-detuned from any electric-dipole line, it should be possible to switch to direct pumping to untrapped $|f\rangle$ states. We have not implemented this yet, due to the technical complexity of doing so.

The Josephson heat interferometer

Francesco Giazotto¹ & María José Martínez-Pérez¹

The Josephson effect¹ is perhaps the prototypical manifestation of macroscopic phase coherence, and forms the basis of a widely used electronic interferometer—the superconducting quantum interference device² (SQUID). In 1965, Maki and Griffin predicted³ that the thermal current through a temperature-biased Josephson tunnel junction coupling two superconductors should be a stationary periodic function of the quantum phase difference between the superconductors: a temperature-biased SQUID should therefore allow heat currents to interfere^{4,5}, resulting in a thermal version of the electric Josephson interferometer. This phase-dependent mechanism of thermal transport has been the subject of much discussion^{4,6–8} but, surprisingly, has yet to be realized experimentally. Here we investigate heat exchange between two normal metal electrodes kept at different temperatures and tunnel-coupled to each other through a thermal ‘modulator’ (ref. 5) in the form of a direct-current SQUID. We find that heat transport in the system is phase dependent, in agreement with the original prediction. Our Josephson heat interferometer yields magnetic-flux-dependent temperature oscillations of up to 21 millikelvin in amplitude, and provides a flux-to-temperature transfer coefficient exceeding 60 millikelvin per flux quantum at 235 millikelvin. In addition to confirming the existence of a phase-dependent thermal current unique to Josephson junctions, our results point the way towards the phase-coherent manipulation of heat in solid-state nanocircuits.

To realize a Josephson heat interferometer we consider a symmetric direct-current SQUID (that is, a superconducting loop comprising two equal Josephson tunnel junctions with resistance R_J) composed of two identical superconductors S_1 and S_2 in thermal steady state and residing at temperatures T_1 and T_2 , respectively (see Fig. 1a). For definiteness, we assume $T_1 \geq T_2$ so that the SQUID is temperature-biased only. With this assumption, the total heat flow \dot{Q}_{SQUID} from S_1 to S_2 becomes stationary and is given by^{3,5,7,8}

$$\dot{Q}_{\text{SQUID}}(\Phi) = 2\dot{Q}_{\text{qp}} - 2\dot{Q}_{\text{int}} \left| \cos\left(\frac{\pi\Phi}{\Phi_0}\right) \right| \quad (1)$$

where $\Phi_0 \approx 2 \times 10^{-15}$ Wb is the flux quantum, the factor 2 accounts for two identical SQUID junctions, and Φ is the applied magnetic flux threading the loop. Φ -dependence appears in equation (1) only through the cosine term so that \dot{Q}_{SQUID} consists of a Φ_0 -periodic function superimposed on a magnetic-flux-independent component. In the above expression

$$\dot{Q}_{\text{qp}}(T_1, T_2) = \frac{2}{e^2 R_J} \int_0^\infty \mathcal{E} \mathcal{N}_1(\mathcal{E}, T_1) \mathcal{N}_2(\mathcal{E}, T_2) [f_1(\mathcal{E}, T_1) - f_2(\mathcal{E}, T_2)] d\mathcal{E}$$

is the usual quasiparticle heat current^{3,9}, whereas $\dot{Q}_{\text{int}}(T_1, T_2) = \frac{2}{e^2 R_J} \int_0^\infty \mathcal{E} \mathcal{M}_1(\mathcal{E}, T_1) \mathcal{M}_2(\mathcal{E}, T_2) [f_1(\mathcal{E}, T_1) - f_2(\mathcal{E}, T_2)] d\mathcal{E}$ is the power flow due to interference between quasiparticles and the Cooper-pairs condensate^{3,4,6–8}. $f_i(\mathcal{E}, T_i) = [1 + \exp(\mathcal{E}/k_B T_i)]^{-1}$ is the Fermi–Dirac distribution at temperature T_i ($i = 1, 2$), $\mathcal{N}_i(\mathcal{E}, T_i) = |\mathcal{E}| / \sqrt{\mathcal{E}^2 - \Delta_i(T_i)^2} \Theta[\mathcal{E}^2 - \Delta_i(T_i)^2]$ is the Bardeen–Cooper–Schrieffer normalized density of states in the superconductors¹⁰, $\mathcal{M}_i(\mathcal{E}, T_i) = \Delta_i(T_i) / \sqrt{\mathcal{E}^2 - \Delta_i(T_i)^2} \Theta[\mathcal{E}^2 - \Delta_i(T_i)^2]$ (refs 3, 4, 6–8), $\Delta_i(T_i)$ is the

temperature-dependent energy gap¹⁰, $\Theta(x)$ is the Heaviside function, k_B is the Boltzmann constant, \mathcal{E} is the energy relative to the chemical potential in the superconductors and e is the charge of the electron. We note that both \dot{Q}_{qp} and \dot{Q}_{int} vanish for $T_1 = T_2$, whereas \dot{Q}_{int} also vanishes when at least one of the superconductors is in the normal state.

The implementation of our heat interferometer is shown in Fig. 1b. The structure has been fabricated with electron-beam lithography and three-angle shadow-mask evaporation. It consists of source and drain copper (Cu) electrodes tunnel-coupled to a superconducting aluminium (Al) island (S_1), defining one branch of a direct-current SQUID. The normal-state resistances of the source and drain junctions are $R_{\text{source}} \approx 1.5$ k Ω and $R_{\text{drain}} \approx 1$ k Ω , respectively, whereas the resistance of each SQUID junction is $R_J \approx 1.3$ k Ω . S_1 is also contacted by an extra Al probe (S_3) via a tunnel junction with normal-state resistance $R_{\text{probe}} \approx 0.55$ k Ω . This tunnel junction was designed so to have a critical current larger than that of the SQUID, thereby allowing an exact determination of the interferometer critical current. Moreover, the contact with S_3 was explicitly made tunnel-like so to limit heat leakage out of S_1 . Both source and drain are tunnel-coupled to a few external Al probes (vertical wires in Fig. 1b) so to realize normal metal–insulator–superconductor (NIS) junctions, with normal-state resistance of about 25 k Ω each, which allow Joule heating and thermometry⁹.

Below the critical temperature of Al (about 1.4 K) Josephson coupling allows dissipation-free charge transport through the SQUID. The SQUID voltage–current characteristics at 240 mK for two representative magnetic-flux values are shown in Fig. 1c. In particular, a well-defined Josephson current with maximum amplitude of $I_c(0) \approx 226$ nA is observed at $\Phi = 0$. Finite-bias switching steps appearing at $\Phi_0/2$ represent the critical current of the S_3S_1 Josephson junction when the SQUID is driven into the dissipative regime¹¹. The magnetic-flux pattern of the SQUID total critical current $I_c(\Phi)$ along with the theoretical prediction^{2,10} is displayed in Fig. 1d, and shows a nearly complete supercurrent modulation which confirms the good symmetry of the SQUID. The SQUID screening parameter is $\beta_L = I_c(0)\mathcal{L}/\Phi_0 \approx 1.6 \times 10^{-3}$ where $\mathcal{L} \approx 15$ pH is the estimated geometric inductance of the loop.

Thermal transport, and hence heat interference in the structure, arises from heating electrons in the source above the lattice temperature (which is the bath temperature, T_{bath}) so as to elevate the quasiparticle temperature in S_1 (T_1) and create a temperature gradient across the SQUID. This hypothesis is expected to hold because the second branch of the SQUID (S_2) extends into a large-volume lead, providing efficient thermalization of its quasiparticles at T_{bath} . \dot{Q}_{SQUID} will thus appear, leading to a Φ_0 -periodic modulation of drain electron temperature (T_{drain}).

Investigation of heat transport in our system is performed as follows. One pair of NIS junctions in the source is operated as a heater, and a second pair of NIS junctions is used to measure electron temperature (T_{source}) in the source by applying a small direct-current bias current and recording the corresponding temperature-dependent voltage drop V_{th} (refs 9, 12). Analogously, another pair of NIS junctions is used to perform thermometry in the drain (see Fig. 1b). Thermometer bias currents were optimized to achieve high sensitivity while limiting the impact of self-heating or self-cooling⁹. Figure 1e

¹NEST, Istituto Nanoscienze—CNR and Scuola Normale Superiore, Piazza San Silvestro 12, I-56127 Pisa, Italy.

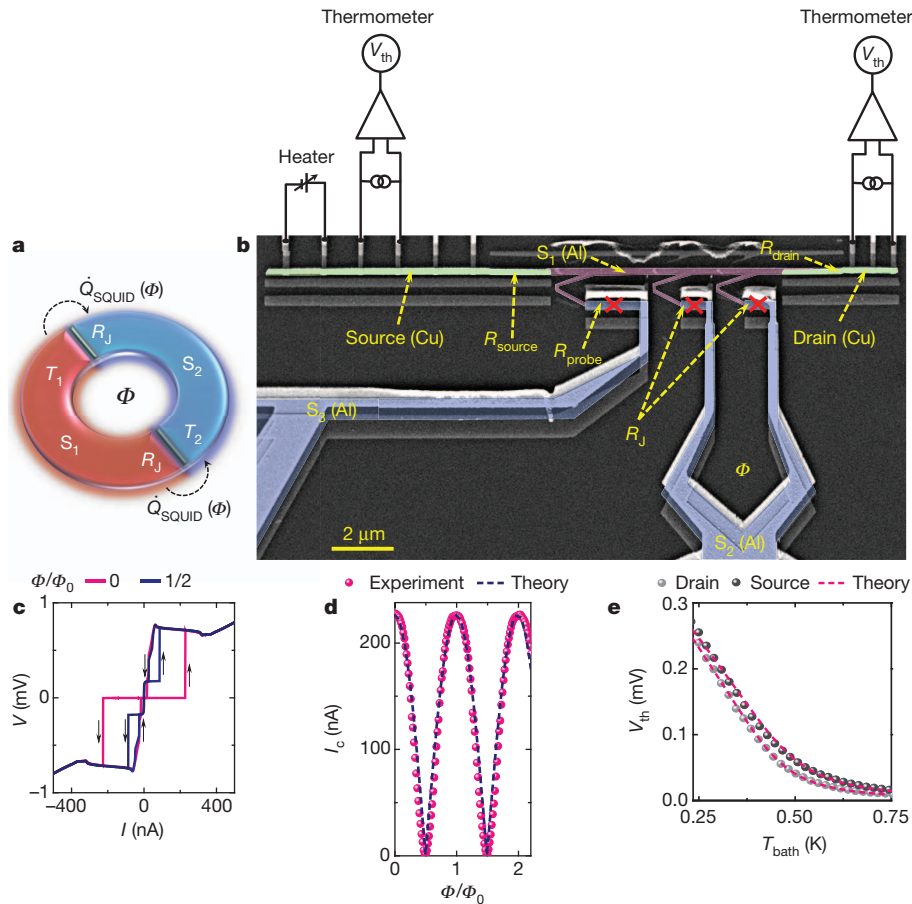


Figure 1 | Josephson heat interferometer. **a**, Scheme of the device core: two identical superconductors S_1 and S_2 kept at different temperatures T_1 and T_2 (with $T_1 \geq T_2$), respectively, are tunnel coupled so as to implement a direct-current SQUID. R_j is the normal-state resistance of each junction, Φ is the applied magnetic flux threading the loop, whereas $\dot{Q}_{\text{SQUID}}(\Phi)$ is the heat current flowing from the hotter to the colder superconductor. **b**, Pseudo-colour scanning electron micrograph of the heat interferometer. Cu source and drain electrodes share the contact through AlO_x tunnel barriers with an Al island (S_1) defining one branch of a direct-current SQUID. The other branch of the SQUID (S_2) extends into a large-volume lead to ensure proper thermalization of its quasiparticles at the bath temperature T_{bath} . S_1 is also in contact with an Al tunnel probe (S_3), enabling independent characterization of the SQUID. S_3 as well as the SQUID Josephson junctions are marked by red crosses. NIS

displays the calibration curves of source and drain thermometers versus T_{bath} , obtained by slowly sweeping the cryostat temperature from 235 mK to 750 mK. The corresponding theoretical results for a NIS junction¹⁰ are also shown.

Measurement of heat interference is done by stabilizing the cryostat temperature at a desired T_{bath} and heating the source up to a given T_{source} . Typical Joule heating (\dot{Q}_{heat}) was in the range 1–35 pW. T_{drain} is then recorded against a slowly sweeping external magnetic flux. Figure 2a shows T_{drain} against Φ measured at 235 mK for increasing values of T_{source} . Notably, T_{drain} is Φ_0 -periodic in Φ , like the Josephson critical current (see Fig. 1d).

As we shall argue, such a temperature modulation is of coherent nature, and stems from the magnetic-flux control of \dot{Q}_{SQUID} , which is a hallmark of the Josephson effect. Raising T_{source} leads to a monotonic enhancement of the average drain temperature over one flux quantum, $\langle T_{\text{drain}} \rangle$, which follows from increased heat flow across the structure. On the other hand, the modulation amplitude δT_{drain} , defined as the difference between the maximum and minimum values of T_{drain} , turns out to increase initially and then tend towards saturation at larger T_{source} . In particular, a δT_{drain} up to about 21 mK is observed,

corresponding to around 9% of relative modulation amplitude at 235 mK. A T_{drain} modulation of amplitude 1.5–2 mK is observed also without intentional source heating (that is, at $T_{\text{source}} = 235$ mK), and might be related to a parasitic power input in the structure (of 1–5 fW) due to coupling to the environment or via the electrical leads, yielding a small quasiparticle overheating localized predominantly in the S_1 electrode.

The full T_{source} -dependence of $\langle T_{\text{drain}} \rangle$ and δT_{drain} are displayed in Fig. 2b and confirm the behaviour described above. We stress that heat interference manifests itself only owing to the existence of a finite temperature bias across the SQUID. Any voltage drop V_{SQUID} occurring at the SQUID Josephson junctions makes the phase-coherent component of \dot{Q}_{SQUID} time-dependent and oscillating at the Josephson frequency, V_{SQUID}/Φ_0 ; it therefore does not contribute to time-averaged direct-current heat transport^{3,4,6}. We also verified that any V_{SQUID} intentionally created by driving the SQUID into the dissipative regime leads to disappearance of heat interference, thus corroborating this picture.

A relevant figure of merit of the heat interferometer is represented by the flux-to-temperature transfer coefficient⁵, $\mathcal{T} \equiv \partial T_{\text{drain}} / \partial \Phi$, shown in Fig. 2c versus Φ for a few selected T_{source} values. It turns

out that \mathcal{T} is Φ_0 -periodic in Φ , with a maximum value of about 0.02 mK/Φ₀ at $T_{\text{source}} = 235$ mK. The modulation amplitude $\delta \mathcal{T}$ is also Φ_0 -periodic in Φ , with a maximum value of about 0.005 mK/Φ₀ at $T_{\text{source}} = 235$ mK. The modulation amplitude $\delta \mathcal{T}$ is also Φ_0 -periodic in Φ , with a maximum value of about 0.005 mK/Φ₀ at $T_{\text{source}} = 235$ mK.

As we shall argue, such a temperature modulation is of coherent nature, and stems from the magnetic-flux control of \dot{Q}_{SQUID} , which is a hallmark of the Josephson effect. Raising T_{source} leads to a monotonic enhancement of the average drain temperature over one flux quantum, $\langle T_{\text{drain}} \rangle$, which follows from increased heat flow across the structure. On the other hand, the modulation amplitude δT_{drain} , defined as the difference between the maximum and minimum values of T_{drain} , turns out to increase initially and then tend towards saturation at larger T_{source} . In particular, a δT_{drain} up to about 21 mK is observed,

out that $|T|$ exceeding $60 \text{ mK } \Phi_0^{-1}$ is obtained at 675 mK . Larger values might be obtained by lowering T_{bath} and by further optimizing the structure design.

To account for our observations we have elaborated a thermal model, which is sketched in Fig. 2d. We assume electrons in S_1 to exchange heat at power \dot{Q}_{source} and \dot{Q}_{drain} owing to quasiparticle heat conduction with the source and drain, respectively, at power \dot{Q}_{SQUID} with S_2 and \dot{Q}_{probe} with S_3 . Quasiparticles in both S_2 and S_3 are assumed to be thermalized at T_{bath} . Furthermore, drain electrons exchange energy at power \dot{Q}_{drain} with S_1 , and at power $\dot{Q}_{\text{e-p,drain}}$ (where subscript ‘e-p’ means ‘electron–phonon’) with lattice phonons residing at T_{bath}

(refs 9, 13). Similarly, source electrons are heated by \dot{Q}_{heat} at T_{source} , and exchange energy at power \dot{Q}_{source} with S_1 and at power $\dot{Q}_{\text{e-p,source}}$ with phonons at T_{bath} . Here, we assume phonons in the various metallic parts of the structure to be well thermalized with substrate phonons residing at cryostat temperature T_{bath} , thereby neglecting lattice heating, given that the Kapitza thermal resistance is negligibly small at such low temperatures¹⁴. The thermal steady state of the S_1 and drain sub-systems may be described by the energy-balance equations

$$\begin{aligned} -\dot{Q}_{\text{source}} + \dot{Q}_{\text{probe}} + \dot{Q}_{\text{SQUID}}(\Phi) + \dot{Q}_{\text{drain}} &= 0 \\ -\dot{Q}_{\text{drain}} + \dot{Q}_{\text{e-p,drain}} &= 0 \end{aligned} \quad (2)$$

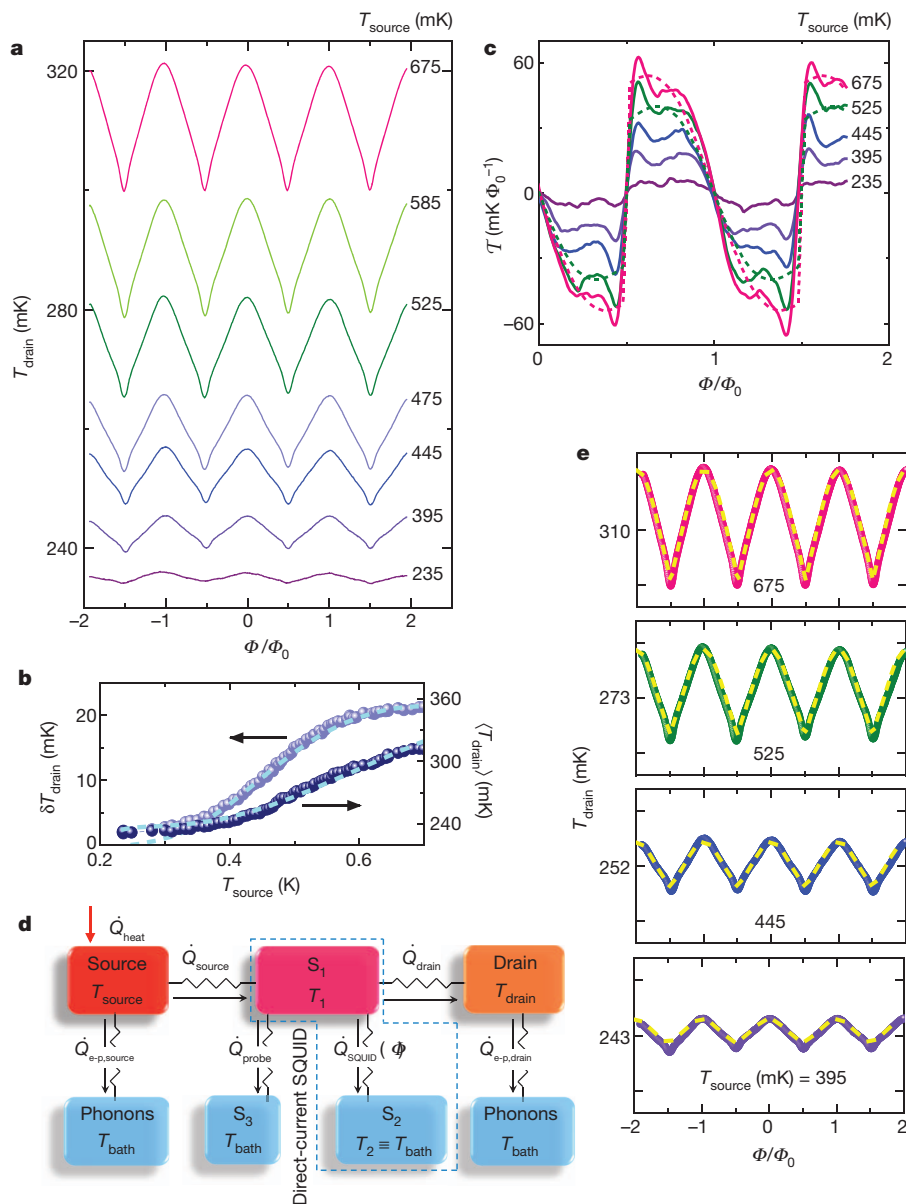


Figure 2 | Behaviour of the heat interferometer at 235 mK. **a**, Flux modulation of T_{drain} measured for several T_{source} values. **b**, Modulation amplitude δT_{drain} (left axis) and average temperature $\langle T_{\text{drain}} \rangle$ (right axis) versus T_{source} . **c**, Flux-to-temperature transfer function $T \equiv \partial T_{\text{drain}} / \partial \Phi$ versus Φ measured at a few selected values of T_{source} . The dashed lines in panels **b** and **c** are the results from the thermal model (see discussion below and in the text). **d**, Idealized thermal diagram accounting for our set-up. Electrons in S_1 exchange energy at power \dot{Q}_{source} and \dot{Q}_{drain} owing to quasiparticle heat conduction with the source and drain, respectively, at power \dot{Q}_{SQUID} with S_2 and \dot{Q}_{probe} with S_3 . Quasiparticles in S_2 and S_3 are assumed to reside at T_{bath} . Drain electrons exchange energy at power \dot{Q}_{drain} with quasiparticles in S_1 , and

at power $\dot{Q}_{\text{e-p,drain}}$ with lattice phonons thermalized at bath temperature. Analogously, source electrons are heated by \dot{Q}_{heat} at T_{source} , and are thermally coupled to S_1 via \dot{Q}_{source} and to the electrode phonons at T_{bath} via $\dot{Q}_{\text{e-p,source}}$. Owing to negligible Kapitza resistance, phonon temperature in the various parts of the structure is assumed to coincide with that of substrate phonons residing at the cryostat temperature T_{bath} . Arrows indicate the direction of heat flows for $T_{\text{bath}} < T_{\text{drain}} < T_1 < T_{\text{source}}$. **e**, Experimental $T_{\text{drain}}(\Phi)$ curves at a few selected values of T_{source} along with the results from thermal model (dashed lines). The full temperature range in each panel is 28 mK, and the vertical division corresponds to 10 mK.

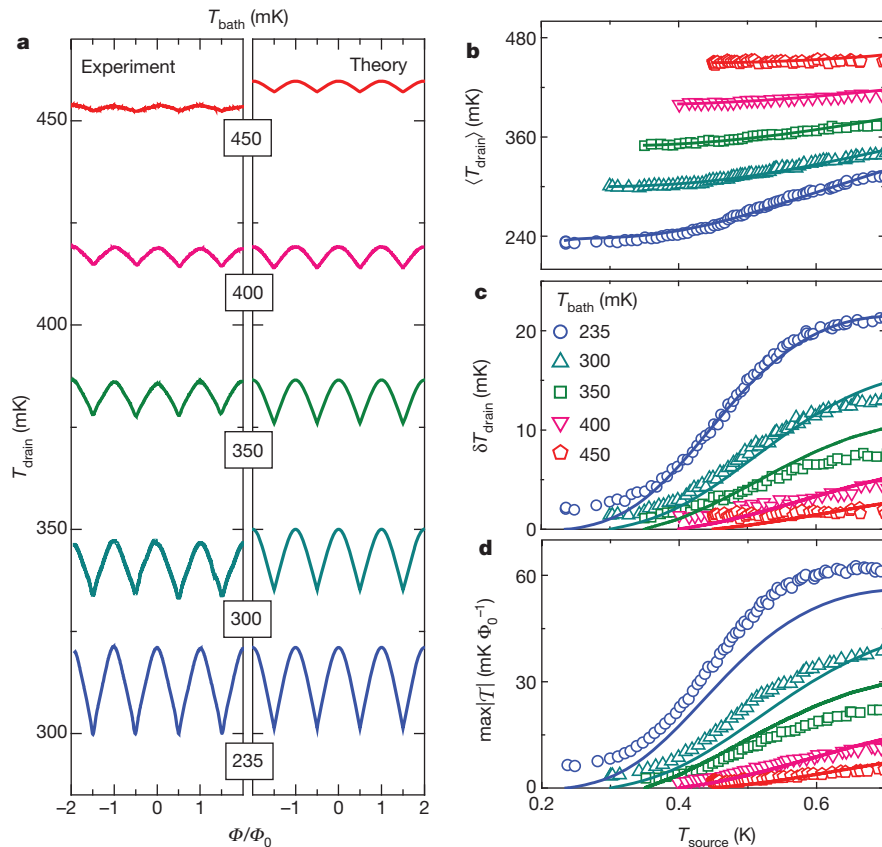


Figure 3 | Behaviour of the heat interferometer at different bath temperatures. **a**, Flux modulation of T_{drain} recorded at different T_{bath} . From bottom to top, data correspond to $T_{\text{source}} = 675$ mK, 700 mK, 690 mK, 700 mK and 700 mK. The left panel shows the experimental data, and the right panel displays results from the thermal model. A sizable temperature modulation is

where the first equation accounts for the thermal budget in S_1 , and the second equation describes heat exchange in the drain. T_1 and T_{drain} can be determined under given conditions for any T_{source} by numerically solving equations (2) (see Methods Summary for further details). The model neglects heat exchange between electrons and photons owing to mismatched impedance^{15–19}, electron–phonon coupling in S_1 due to its reduced volume and low experimental T_{bath} (ref. 20), and phonon heat current³.

Figure 2e shows the comparison between experiments and the model, displaying a few $T_{\text{drain}}(\Phi)$ traces (solid lines) along with the theoretical behaviour (dashed lines). Analogously, dashed lines in Fig. 2b show the predicted $\langle T_{\text{drain}} \rangle$ and δT_{drain} against T_{source} , and in Fig. 2c the predicted $T(\Phi)$ for the two highest T_{source} values. Although idealized, the model provides reasonable agreement with our observations, and grasps the relevant physical picture at the origin of heat interference in our system.

The dependence on bath temperature is shown in Fig. 3a (experimental values, on the left) which displays $T_{\text{drain}}(\Phi)$ at several increasing T_{bath} values for T_{source} set at approximately 700 mK (see Fig. 3 legend). The theoretical values on the right of Fig. 3a show the curves obtained from the model at the corresponding T_{bath} . Besides leading to a monotonic enhancement of $\langle T_{\text{drain}} \rangle$, increasing T_{bath} suppresses δT_{drain} and smears $T_{\text{drain}}(\Phi)$ (ref. 5), which can be mainly ascribed to the enhancement of electron–phonon coupling in the drain⁹ as well as to the influence of thermal broadening. $\delta T_{\text{drain}} \approx 2.5$ mK is still observable at 450 mK, whereas the modulation disappears for $T_{\text{bath}} \gtrsim 500$ mK. We emphasize that $T_{\text{bath}} \gtrsim 500$ mK is substantially smaller than the temperature (about 1.4 K) at which the Josephson effect disappears in the SQUID. Figure 3b–d shows $\langle T_{\text{drain}} \rangle$, δT_{drain} , and the maximum of $|T|$

still observable at 450 mK, whereas δT_{drain} vanishes for $T_{\text{bath}} \geq 500$ mK. **b**, Average temperature $\langle T_{\text{drain}} \rangle$ versus T_{source} . **c**, Modulation amplitude δT_{drain} versus T_{source} . **d**, Maximum value of $|T|$ versus T_{source} . Data in **b**–**d** were measured at the same T_{bath} as in **a**. Solid lines correspond to the thermal model.

versus T_{source} , respectively, recorded at the same T_{bath} as in Fig. 3a, along with the lines obtained from the thermal model. Each of these quantities displays a similar qualitative behaviour at different T_{bath} , and a smoothing of their characteristics is observed by increasing T_{bath} . This is consistent with the picture provided by the model, which captures the main features of the experimental data.

Our results confirm the prediction of almost fifty years ago³, that is, the existence of a phase-dependent component in the heat current flowing through a temperature-biased Josephson tunnel junction. Besides offering insight into thermal transport in Josephson weak links, this effect could represent a valuable tool for phase-coherent manipulation of heat in solid-state nanocircuits^{9,21–25}. We envision novel-concept coherent caloritronic devices, such as heat transistors and thermal splitters, that would exploit phase-dependent heat transfer peculiar to the Josephson effect.

METHODS SUMMARY

The structures were fabricated with electron-beam lithography and three-angle shadow-mask evaporation of metals onto an oxidized Si wafer through a suspended resist mask. In the electron-beam evaporator, the chip was initially tilted at an angle of 28° , and 20 nm of Al was deposited to form S_2 and S_3 . The sample was then exposed to 380 mTorr of O_2 for 4.5 min to form the SQUID tunnel barriers, after which it was tilted to 0° for the deposition of 25 nm of Al to form S_1 , the heaters and the thermometer probes. The chip was subsequently exposed to 800 mTorr of O_2 for 4.5 min to form the heaters, thermometers, and the source and drain tunnel junctions. Finally, 30 nm of Cu was deposited at 42° to form the source and drain.

The magneto-electric characterization of the samples was performed down to 235 mK in a filtered ^3He cryostat. Current biasing of the thermometers was obtained through battery-powered floating sources, whereas voltage and current

were measured with room-temperature preamplifiers. Flux-to-temperature transfer functions were measured using a low-frequency lock-in technique by superimposing a small modulation to the applied magnetic field.

In the energy-balance equations (see equations (2)), $\dot{Q}_{\text{probe}} = \dot{Q}_{\text{qp}}^{\text{probe}} - \dot{Q}_{\text{int}}^{\text{probe}}$ (refs 3–8), $\dot{Q}_{\text{qp}}^{\text{probe}} = \frac{2}{e^2 R_{\text{probe}}} \int_0^\infty \varepsilon \mathcal{N}_1(\varepsilon) \mathcal{N}_3(\varepsilon) [f_1(\varepsilon) - f_3(\varepsilon)] d\varepsilon$, $\dot{Q}_{\text{int}}^{\text{probe}} = \frac{2}{e^2 R_{\text{probe}}} \int_0^\infty \varepsilon \mathcal{M}_1(\varepsilon) \mathcal{M}_3(\varepsilon) [f_1(\varepsilon) - f_3(\varepsilon)] d\varepsilon$, $\mathcal{N}_3(\varepsilon) = \mathcal{N}_2(\varepsilon)$, $\mathcal{M}_3(\varepsilon) = \mathcal{M}_2(\varepsilon)$ and $f_3(\varepsilon) = f_2(\varepsilon)$. Furthermore, $\dot{Q}_{\text{source}} = \frac{2}{e^2 R_{\text{source}}} \int_0^\infty \varepsilon \mathcal{N}_1(\varepsilon) [f_{\text{source}}(\varepsilon) - f_1(\varepsilon)] d\varepsilon$, $\dot{Q}_{\text{drain}} = \frac{2}{e^2 R_{\text{drain}}} \int_0^\infty \varepsilon \mathcal{N}_1(\varepsilon) [f_1(\varepsilon) - f_{\text{drain}}(\varepsilon)] d\varepsilon$ (ref. 9), $f_{\text{source(drain)}}(\varepsilon) = [1 + \exp(\varepsilon/k_B T_{\text{source(drain)}})]^{-1}$, $\dot{Q}_{\text{e-p,drain}} = \Sigma V_{\text{drain}} (T_{\text{drain}}^5 - T_{\text{bath}}^5)$ (refs 9, 13), $\Sigma = 3 \times 10^9 \text{ W K}^{-5} \text{ m}^{-3}$ is the electron-phonon coupling constant for our Cu electrodes⁹, and $V_{\text{drain}} = 2 \times 10^{-20} \text{ m}^3$ is the drain volume. For the solution of equations (2) we additionally set $R_{\text{source}} = 1.5 \text{ k}\Omega$, $R_{\text{drain}} = 1.0 \text{ k}\Omega$, $R_I = 1.3 \text{ k}\Omega$ and $\Delta(0) = 200 \mu\text{eV}$, as determined from the experiment, and varied R_{probe} as the only fitting parameter between 100% and 125% of its nominal value to match data at different values of T_{bath} .

Received 15 May; accepted 23 October 2012.

- Josephson, B. D. Possible new effects in superconductive tunneling. *Phys. Lett.* **1**, 251–253 (1962).
- Clarke, J. & Braginski, A. I. (eds) *The SQUID Handbook* (Wiley-VCH, 2004).
- Maki, K. & Griffin, A. Entropy transport between two superconductors by electron tunneling. *Phys. Rev. Lett.* **15**, 921–923 (1965).
- Guttman, G. D., Ben-Jacob, E. & Bergman, D. J. Interference effect heat conductance in a Josephson junction and its detection in an rf SQUID. *Phys. Rev. B* **57**, 2717–2719 (1998).
- Giazotto, F. & Martínez-Pérez, M. J. Phase-controlled superconducting heat-flux quantum modulator. *Appl. Phys. Lett.* **101**, 102601 (2012).
- Guttman, G. D., Nathanson, B., Ben-Jacob, E. & Bergman, D. J. Phase-dependent thermal transport in Josephson junctions. *Phys. Rev. B* **55**, 3849–3855 (1997).
- Zhao, E., Löfwander, T. & Sauls, J. A. Phase modulated thermal conductance of Josephson weak links. *Phys. Rev. Lett.* **91**, 077003 (2003).
- Zhao, E., Löfwander, T. & Sauls, J. A. Heat transport through Josephson point contacts. *Phys. Rev. B* **69**, 134503 (2004).
- Giazotto, F., Heikkilä, T. T., Luukanen, A., Savin, A. M. & Pekola, J. P. Opportunities for mesoscopes in thermometry and refrigeration: physics and applications. *Rev. Mod. Phys.* **78**, 217–274 (2006).
- Tinkham, M. *Introduction to Superconductivity* 2nd edn (McGraw-Hill, 1996).
- Quaranta, O., Spathis, P., Beltram, F. & Giazotto, F. Cooling electrons from 1 K to 0.4 K with V-based nanorefrigerators. *Appl. Phys. Lett.* **98**, 032501 (2011).
- Nahum, M. & Martinis, J. M. Ultrasensitive-hot-electron microbolometer. *Appl. Phys. Lett.* **63**, 3075–3077 (1993).
- Roukes, M. L., Freeman, M. R., Germain, R. S., Richardson, R. C. & Ketchen, M. B. Hot electrons and energy transport in metals at millikelvin temperatures. *Phys. Rev. Lett.* **55**, 422–425 (1985).
- Wellstood, F. C., Urbina, C. & Clarke, J. Hot-electron effects in metals. *Phys. Rev. B* **49**, 5942–5955 (1994).
- Schmidt, D. R., Schoelkopf, R. J. & Cleland, A. N. Photon-mediated thermal relaxation of electrons in nanostructures. *Phys. Rev. Lett.* **93**, 045901 (2004).
- Meschke, M., Guichard, W. & Pekola, J. P. Single-mode heat conduction by photons. *Nature* **444**, 187–190 (2006).
- Timofeev, A. V., Helle, M., Meschke, M., Möttönen, M. & Pekola, J. P. Electronic refrigeration at the quantum limit. *Phys. Rev. Lett.* **102**, 200801 (2009).
- Peltonen, J. T. et al. Thermal conductance by the inverse proximity effect in a superconductor. *Phys. Rev. Lett.* **105**, 097004 (2010).
- Pascal, L. M. A., Courtois, H. & Hekking, F. W. J. Circuit approach to photonic heat transport. *Phys. Rev. B* **83**, 125113 (2011).
- Timofeev, A. V. et al. Recombination-limited energy relaxation in a Bardeen-Cooper-Schrieffer superconductor. *Phys. Rev. Lett.* **102**, 017003 (2009).
- Dubi, Y. & Di Ventra, M. Heat flow and thermoelectricity in atomic and molecular junctions. *Rev. Mod. Phys.* **83**, 131–155 (2011).
- Ojanen, T. & Jauho, A.-P. Mesoscopic photon heat transistor. *Phys. Rev. Lett.* **100**, 155902 (2008).
- Ruokola, T., Ojanen, T. & Jauho, A.-P. Thermal rectification in nonlinear quantum circuits. *Phys. Rev. B* **79**, 144306 (2009).
- Ryazanov, T. T. & Schmidt, V. V. Observation of thermal electromotive force oscillations versus magnetic vector potential field in superconducting loop with two Josephson SNS junctions. *Solid State Commun.* **42**, 733–735 (1982).
- Panaitov, G. I., Ryazanov, V. V. & Schmidt, V. V. Thermoelectric ac Josephson effect in SNS junctions. *Phys. Lett.* **100**, 301–303 (1984).

Acknowledgements We thank F. Taddei for discussions and for a careful reading of the manuscript. We also thank C. Altimiras, C. W. J. Beenakker, M. Di Ventra, T. T. Heikkilä, M. A. Laakso, F. Portier and P. Spathis for comments, and the EC FP7 programme number 228464 “Microkelvin” for partial financial support.

Author Contributions F.G. conceived and designed the experiment, performed and analysed the measurements, and developed the theoretical model. M.J.M.-P. fabricated the samples, contributed to the measurements, and analysed the data. F.G. wrote the manuscript with input from M.J.M.-P.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.G. (giazotto@sns.it).

Fractionalized excitations in the spin-liquid state of a kagome-lattice antiferromagnet

Tian-Heng Han¹, Joel S. Helton², Shaoyan Chu³, Daniel G. Nocera⁴, Jose A. Rodriguez-Rivera^{2,5}, Collin Broholm^{2,6} & Young S. Lee¹

The experimental realization of quantum spin liquids is a long-sought goal in physics, as they represent new states of matter. Quantum spin liquids cannot be described by the broken symmetries associated with conventional ground states. In fact, the interacting magnetic moments in these systems do not order, but are highly entangled with one another over long ranges¹. Spin liquids have a prominent role in theories describing high-transition-temperature superconductors^{2,3}, and the topological properties of these states may have applications in quantum information⁴. A key feature of spin liquids is that they support exotic spin excitations carrying fractional quantum numbers. However, detailed measurements of these 'fractionalized excitations' have been lacking. Here we report neutron scattering measurements on single-crystal samples of the spin-1/2 kagome-lattice antiferromagnet $\text{ZnCu}_3(\text{OD})_6\text{Cl}_2$ (also called herbertsmithite), which provide striking evidence for this characteristic feature of spin liquids. At low temperatures, we find that the spin excitations form a continuum, in contrast to the conventional spin waves expected in ordered antiferromagnets. The observation of such a continuum is noteworthy because, so far, this signature of fractional spin excitations has been observed only in one-dimensional systems. The results also serve as a hallmark of the quantum spin-liquid state in herbertsmithite.

In a spin liquid, the atomic magnetic moments are strongly correlated but do not order or freeze even in the limit as the temperature, T , goes to zero. Although many types of quantum spin-liquid states exist in theory, a feature that is expected to be common to all is the presence of deconfined spinons as an elementary excitation from the ground state¹. Spinons are spin-half ($S = 1/2$) quantum excitations into which conventional spin-wave excitations with $S = 1$ fractionalize. In one dimension, this phenomenon is well established for the $S = 1/2$ Heisenberg antiferromagnetic chain, where spinons may be thought of as magnetic domain boundaries that disrupt Néel order and are free to propagate away from each other. In the one-dimensional compound KCuF_3 , a continuum of spinon excitations has been well characterized using neutron scattering⁵. In two dimensions, the nature of the spinon excitations is less clear. First, the existence of two-dimensional magnets with a quantum spin-liquid ground state is still a matter of great debate. Second, the various spin-liquid states which are proposed in theory give rise to a variety of spinon excitation spectra, which may be either gapped or gapless.

The $S = 1/2$ kagome-lattice Heisenberg antiferromagnet has long been recognized as a promising system in which to search for quantum spin-liquid states, because the kagome network of corner-sharing triangles frustrates long-range magnetic order^{6–8}. We have devised synthetic methods to produce herbertsmithite ($\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$) in which the $S = 1/2$ Cu^{2+} moments are arranged on a structurally perfect kagome lattice⁹ and nonmagnetic Zn^{2+} ions separate the lattice planes. A depiction of the crystal structure is shown in Supplementary Fig. 1. Whereas herbertsmithite typically contains a small percentage of excess

Cu^{2+} ions ($\sim 5\%$ of the total) substituting for Zn^{2+} ions in the inter-layer sites, the kagome planes contain only Cu^{2+} ions¹⁰. Measurements on powder samples^{11–13} indicate strong antiferromagnetic superexchange ($J \approx 17$ meV, where J is the exchange coupling that appears in the nearest-neighbour Heisenberg Hamiltonian) and the absence of long-range magnetic order or spin freezing down to temperatures of $T = 0.05$ K. The bulk magnetic properties reveal a small Dzyaloshinskii–Moriya interaction and an easy-axis exchange anisotropy^{14,15}, both of order $J/10$. Despite these small imperfections, the nearest-neighbour Heisenberg model on a kagome lattice is still an excellent approximation of the spin Hamiltonian for herbertsmithite. This is especially important, because recent calculations on record lattice sizes indicate that the ground state of this model is in fact a quantum spin liquid¹⁶. Thus, experiments to probe the spin correlations in herbertsmithite are all the more urgent.

To this end, we recently succeeded in developing a technique for the growth of large, high-quality single crystals of herbertsmithite¹⁷, and small pieces have been used in studies involving local probes^{18,19}, anomalous X-ray diffraction¹⁰, susceptibility¹⁵ and Raman scattering²⁰. In this Letter, we report inelastic neutron scattering measurements on a large, deuterated, single-crystal sample of herbertsmithite. The neutron scattering cross-section is directly proportional to the dynamic structure factor $S_{\text{tot}}(\mathbf{Q}, \omega)$ (where \mathbf{Q} and ω stand for the momentum and energy transferred to the sample, respectively), which includes both the nuclear and magnetic signals. The magnetic part, $S_{\text{mag}}(\mathbf{Q}, \omega)$, is the Fourier transform (in time and space) of the spin–spin correlation function and can be obtained by subtracting the nuclear scattering as described in the Supplementary Information. After calibration with respect to a vanadium standard, the measured structure factors are expressed in absolute units.

Contour plots of $S_{\text{tot}}(\mathbf{Q}, \omega)$ are shown in Fig. 1a–c for $T = 1.6$ K and three different energy transfers $\hbar\omega$ (\hbar denotes Planck's constant divided by 2π). Figure 1a shows data for $\hbar\omega = 6$ meV. Surprisingly, the scattered intensity is exceedingly diffuse, spanning a large fraction of the hexagonal Brillouin zone. A similar pattern of diffuse scattering is observed for $\hbar\omega = 2$ meV (Fig. 1b). The diffuse nature of the scattering at a temperature that is two orders of magnitude below the exchange energy scale, J , is in strong contrast to observations in non-frustrated quantum magnets. The $S = 1/2$ square-lattice antiferromagnet La_2CuO_4 develops substantial antiferromagnetic correlations for $T < J/2$ (ref. 21), temperatures at which the low-energy scattering is strongly peaked near the (π, π) point in reciprocal space. In herbertsmithite, the scattered intensity is not strongly peaked at any specific point, and this remains true for all energies measured from $\hbar\omega = 0.25$ to 11 meV. This behaviour is also markedly different from that observed in the larger, $S = 5/2$ kagome antiferromagnet $\text{KFe}_3(\text{OH})_6(\text{SO}_4)_2$ which becomes magnetically ordered at low temperatures and has magnetic peaks at $q = 0$ wavevectors above the ordering temperature²².

¹Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ²NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA. ³Center for Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁴Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁵Department of Materials Science and Engineering, University of Maryland, College Park, Maryland 20742, USA. ⁶Institute for Quantum Matter and Department of Physics and Astronomy, The Johns Hopkins University, Baltimore, Maryland 21218, USA.

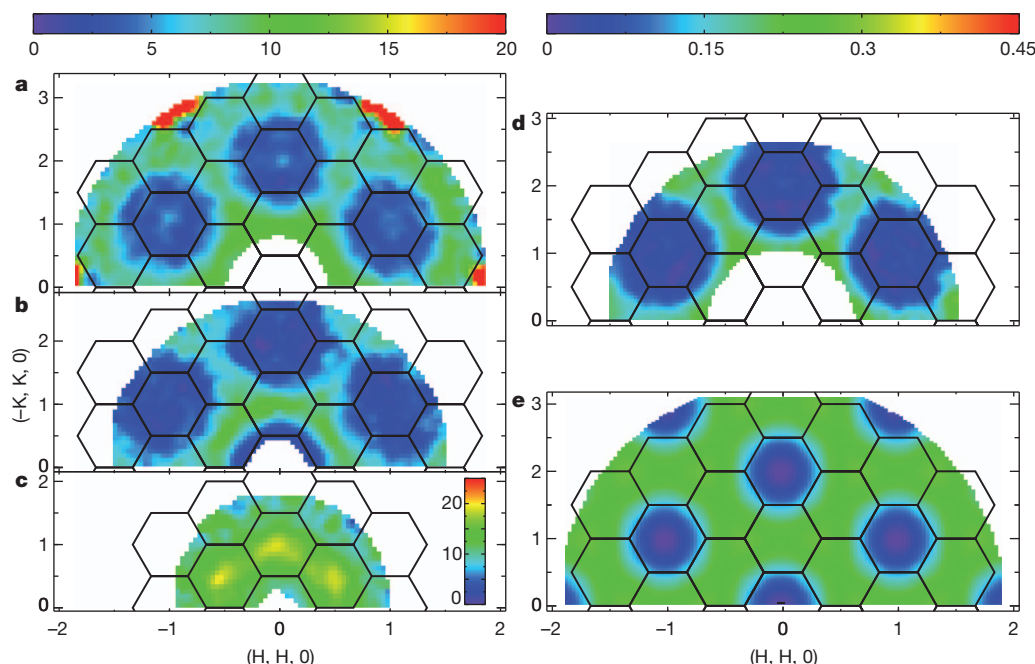


Figure 1 | Inelastic neutron scattering from the spin excitations, plotted in reciprocal space. **a–c**, Measurements were made at $T = 1.6$ K on a single-crystal sample of $\text{ZnCu}_3(\text{OD})_6\text{Cl}_2$. The dynamic structure factor, $S_{\text{tot}}(\mathbf{Q}, \omega)$, is plotted for $\hbar\omega = 6$ meV (**a**) and $\hbar\omega = 2$ meV (**b**) with $E_f = 5.1$ meV and $\hbar\omega = 0.75$ meV (**c**) with $E_f = 3.0$ meV. The background was measured with an empty sample holder and subtracted. The diffuse scattering is mostly magnetic in origin, because the phonon contribution to the signal is small (except near the $(2, 2, 0)$ -type positions, where the fundamental Bragg peaks are strong). **d**, The magnetic part of the dynamic structure factor, $S_{\text{mag}}(\mathbf{Q}, \omega)$, integrated

over $1 \leq \hbar\omega \leq 9$ meV. **e**, Calculation of the equal-time structure factor, $S_{\text{mag}}(\mathbf{Q})$, for a model of uncorrelated nearest-neighbour dimers. The intensity corresponds to $1/8$ of the total moment sum rule $S(S+1)$ for the spins on the kagome lattice. The data presented in **a–c** are expressed in barn $\text{sr}^{-1} \text{eV}^{-1}$ per formula unit, as shown by the left colour bars. The data presented in parts **d** and **e** are dimensionless, with the scale given by the right colour bar. The Brillouin zone boundaries are drawn in the figure for clarity; they correspond to the conventional unit cell with parameters $a = b = 6.83$ Å, $c = 14.05$ Å, $\alpha = \beta = 90^\circ$ and $\gamma = 120^\circ$.

The observed \mathbf{Q} dependence of the scattered intensity provides important information on the ground-state spin correlations. The scattering in reciprocal space has the shape of broadened hexagonal rings centred at $(0, 0, 0)$ - and $(2, 0, 0)$ -type positions. All of the scans that we have performed from $\hbar\omega = 1.5$ to 11 meV show similar patterns for the scattered magnetic intensity. The energy-integrated dynamic structure factor over the integration range $1 \leq \hbar\omega \leq 9$ meV is plotted in Fig. 1d. This quantity serves as an approximation of the equal-time structure factor. For comparison, a calculation of the equal-time structure factor for a collection of uncorrelated nearest-neighbour singlets on a kagome lattice is shown in Fig. 1e. To a first approximation, the observed magnetic signal resembles this calculation. Therefore, the ground-state wavefunction of herbertsmithite has a large component resembling randomly arranged nearest-neighbour singlets, similar to a short-range resonating valence-bond state^{21,6,23}. However, it is also clear that the data have a narrower width in reciprocal space than does the model calculation. Thus, the spin–spin correlations in herbertsmithite extend beyond nearest neighbours, as further discussed below. The intensity in Fig. 1e corresponds to $1/8$ of the total moment sum rule²⁴. For the data, the integrated intensity up to $\hbar\omega = 11$ meV corresponds to 20(3)% of the total moment (where the uncertainty represents 1 s.d.). This indicates that the excitations extend up to much higher energies (a few multiples of J), and future inelastic measurements up to these energies would be of great interest.

At the lowest measured energy transfers, we observe additional features in the pattern of magnetic scattering. Figure 1c depicts the intensity contour plot for $\hbar\omega = 0.75$ meV, showing additional broad peaks centred at $(1, 0, 0)$ and equivalent positions (seen as the yellowish spots near the centres of the low- \mathbf{Q} Brillouin zones). The $(1, 0, 0)$ position does not correspond to a nuclear Bragg position for this crystal structure. Additional scans taken with $\hbar\omega$ between 0.25 and 1 meV confirm that this feature is generic to the low-energy transfers. This peak is probably influenced by the weakly coupled Cu^{2+} ions at

the interlayer Zn^{2+} sites, which are believed to affect the low-energy scattering²⁵.

The scattering pattern's overall insensitivity to energy transfer is another remarkable feature of the data. Conventional spin-wave excitations take the form of sharp surfaces of dispersion in \mathbf{Q} – ω space. Such spin-wave excitations were indeed observed in the $S = 5/2$ kagome antiferromagnet $\text{KFe}_3(\text{OH})_6(\text{SO}_4)_2$ (ref. 26). In herbertsmithite, no surfaces of dispersion are observable in the low-temperature data. The dependence of $S_{\text{tot}}(\mathbf{Q}, \omega)$ on $\hbar\omega$ and \mathbf{Q} is plotted in Fig. 2 for two high-symmetry directions in reciprocal space: the $(\text{H}, 0, 0)$ direction (Fig. 2a) and the $(\text{H}, \text{H}, 0)$ direction (Fig. 2b). These directions are indicated by thick black lines in Fig. 2d. These plots show that the spin excitations form a broad, continuous band (or a continuum), extending up to the highest measured energy, 11 meV. This is direct evidence that the excitations are fractionalized, forming a continuum in this two-dimensional antiferromagnet.

In Fig. 2c and its inset, the energy dependences of $S_{\text{tot}}(\mathbf{Q}, \omega)$ and $S_{\text{mag}}(\mathbf{Q}, \omega)$ are plotted for high symmetry \mathbf{Q} positions as indicated in the reciprocal space map in Fig. 2d. The scattered signal is rather flat for $2 \leq \hbar\omega \leq 10$ meV but increases significantly with decreasing energy transfer below $\hbar\omega = 1.5$ meV. There is no indication of a spin gap down to $\hbar\omega = 0.25$ meV at the measured reciprocal space positions.

The magnetic intensity can be plotted as one-dimensional 'line scans' along specific directions in reciprocal space. In Fig. 3a, $S_{\text{mag}}(\mathbf{Q}, \omega)$ is shown along the $(-2, 1 + \mathbf{K}, 0)$ direction, indicated by the thick red line on the reciprocal space map in Fig. 3d. Three energy transfers, $\hbar\omega = 2, 6$ and 10 meV, are plotted, and there is no substantial change in the peak width as a function of energy transfer. The width of these line scans, determined by fits, can be found in Supplementary Fig. 3. In Fig. 3b, $S_{\text{mag}}(\mathbf{Q}, \omega)$ is integrated over $1 \leq \hbar\omega \leq 11$ meV and compared with the calculated equal-time $S_{\text{mag}}(\mathbf{Q}, \omega)$ for uncorrelated nearest-neighbour singlets. The solid line corresponds to the result of the uncorrelated nearest-neighbour singlet model multiplied by $|F(\mathbf{Q})|^2$, where $F(\mathbf{Q})$ is

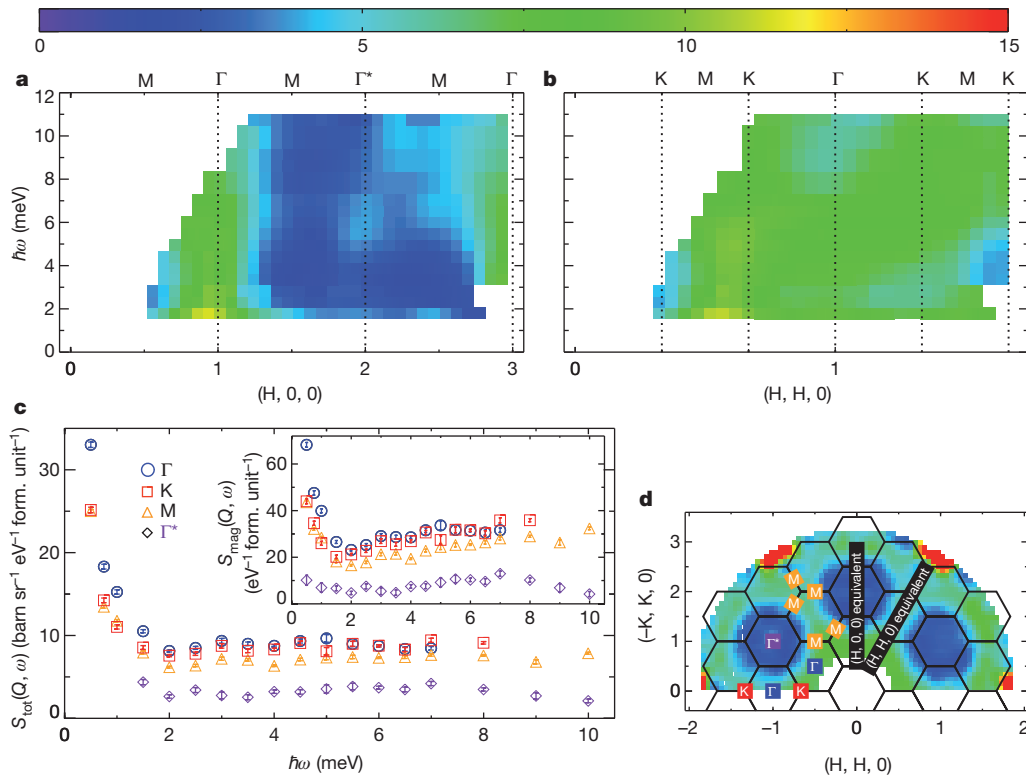


Figure 2 | Inelastic neutron scattering measured along symmetry directions and at high-symmetry locations. **a, b,** Intensity contour plots of the dynamic structure factor as a function of $\hbar\omega$ and Q for the $(H, 0, 0)$ direction (**a**) and the $(H, H, 0)$ direction (**b**). These directions are indicated by the thick black lines on the reciprocal space map shown in **d**. Along the $(H, H, 0)$ direction, a broad excitation continuum is observed over the entire range measured. The colour bar shows the magnitude of $S_{\text{tot}}(Q, \omega)$ in barn $\text{sr}^{-1} \text{eV}^{-1}$ per formula unit. **c,** Energy dependence of $S_{\text{tot}}(Q, \omega)$ measured at high-symmetry reciprocal

space locations. Data for $\hbar\omega \geq 1.5$ meV were measured with $E_f = 5.1$ meV, whereas data for $\hbar\omega \leq 1$ meV were measured with $E_f = 3.0$ meV for better energy resolution (except those at Γ^* , which were measured with $E_f = 5.1$ meV). Error bars, 1 s.d. Inset, energy dependence of $S_{\text{mag}}(Q, \omega)$ with the non-magnetic scattering from the sample subtracted. $S_{\text{mag}}(Q, \omega)$ is normalized to have units of eV^{-1} per formula unit, consistent with the magnetic structure factor defined in Supplementary Information. **d,** The integrated areas in reciprocal space referred to **a–c**.

the free- Cu^{2+} magnetic form factor. Here the measured data indicate longer-range correlations than the nearest-neighbour singlet model. Figure 3c depicts a line scan of the dynamic structure factor (integrated

over $1 \leq \hbar\omega \leq 7$ meV) along the $(0, K, 0)$ direction. The nearest-neighbour singlet model does not account for the observed scattering intensity at the $(0, 2, 0)$ -type positions.

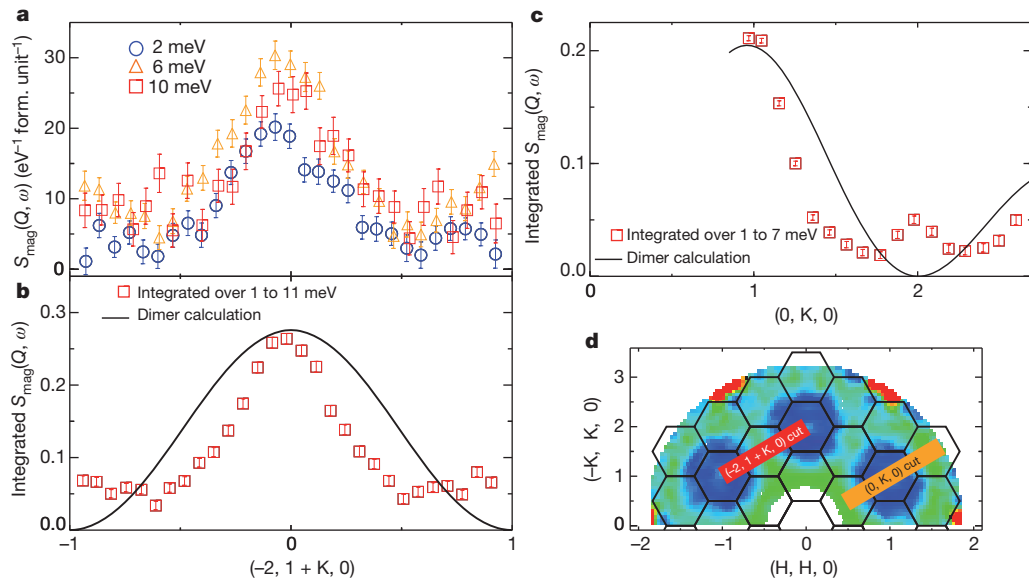


Figure 3 | The measured dynamic structure factor along specific directions in reciprocal space with comparison to the nearest-neighbour singlet model. **a,** $S_{\text{mag}}(Q, \omega)$ along the $(-2, 1 + K, 0)$ direction, indicated by the thick red line on the reciprocal space map in **d**. Three energy transfers, $\hbar\omega = 2, 6$ and 10 meV, are shown. **b,** $S_{\text{mag}}(Q, \omega)$ along the $(-2, 1 + K, 0)$ direction integrated over

$1 \leq \hbar\omega \leq 11$ meV. **c,** $S_{\text{mag}}(Q, \omega)$ along the $(0, K, 0)$ direction, indicated by the thick orange line on the reciprocal space map in **d**, integrated over $1 \leq \hbar\omega \leq 7$ meV. The solid lines in **b** and **c** are the calculated equal-time structure factors for uncorrelated nearest-neighbour singlets multiplied by $|F(Q)|^2$. **d,** The trajectories in reciprocal space referred to **a–c**. Error bars, 1 s.d.

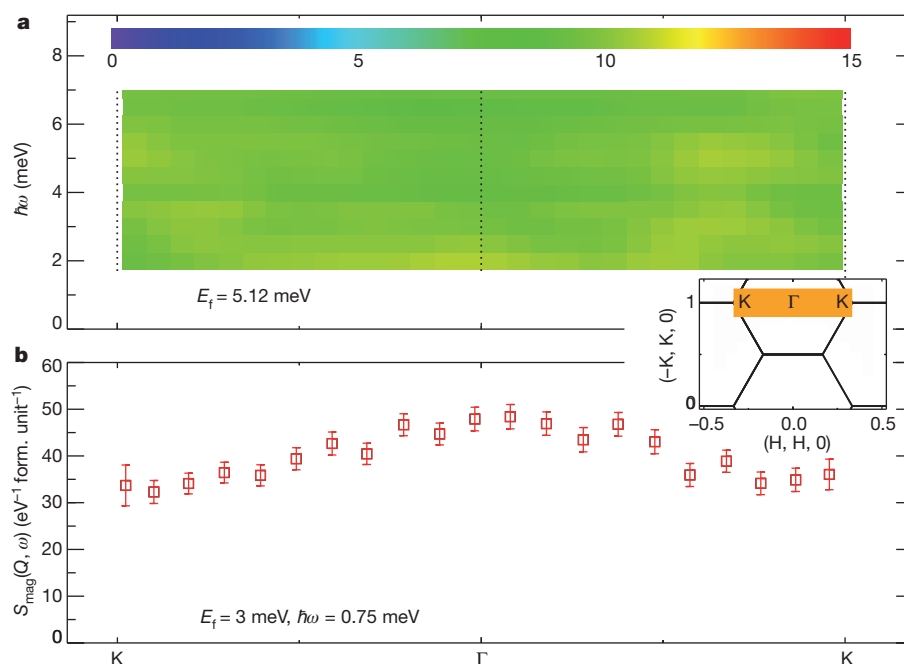


Figure 4 | The measured dynamic structure factor along the K–Γ–K direction within the (1,0,0) Brillouin zone. **a**, Intensity contour plot of the dynamic structure factor as a function of $\hbar\omega$ and Q . Inset, region of integration along K–Γ–K shown as a thick orange line. An excitation continuum is

observed. The colour bar indicates the magnitude of $S_{\text{tot}}(Q, \omega)$ in barn $\text{sr}^{-1} \text{eV}^{-1}$ per formula unit. **b**, A line cut plotting $S_{\text{mag}}(Q, \omega)$ along the K–Γ–K direction for $\hbar\omega = 0.75 \text{ meV}$. Error bars, 1 s.d.

Further evidence of the continuum nature of the scattering is shown in Fig. 4a, where $S_{\text{tot}}(Q, \omega)$ is plotted along the K–Γ–K direction in the (1,0,0) Brillouin zone. For $2 \leq \hbar\omega \leq 7 \text{ meV}$, the scattered intensity is nearly constant along this direction. The data also show another point of contrast to the nearest-neighbour singlet calculation, which predicts slightly larger intensities near the K points, which is not seen in the measurements. Instead, at low energy transfers, $\hbar\omega < 2 \text{ meV}$, the intensity has a broad maximum at the Γ point, as shown in Fig. 4b for $\hbar\omega = 0.75 \text{ meV}$. Interestingly, available theoretical calculations for $S_{\text{mag}}(Q, \omega)$ based on spinon excitations^{6,27} are reasonably consistent with the measured intensity pattern, except for discrepancies near the K points.

A central question for classification of the ground state of herbertsmithite is whether a spin gap exists. One surprising aspect of our data is that the spin excitations seem to be gapless over a wide range of Q positions, at least down to $\hbar\omega = 0.25 \text{ meV}$. This observation is difficult to reconcile with the ground-state properties of valence-bond crystals²⁸ or gapped spin liquids (such as a short-range resonating valence-bond state). Even most theories for gapless spin liquids predict only a small set of reciprocal lattice points for which the excitations are truly gapless^{29,30}. One possible caveat to our finding is that the small percentage of weakly interacting impurities in the interlayer sites may hide the intrinsic spin gap of the kagome spins. However, it is likely that the impurities affect only the excitations below 1 meV, where the upturn in intensity is seen with decreasing energy transfer. Thus, the hexagonal ring pattern of the structure factor for $1.5 \leq \hbar\omega \leq 11 \text{ meV}$ is undoubtedly intrinsic to the kagome layers. And, consequently, this sets a conservative upper bound for the intrinsic spin gap of $\sim J/10$, if a gap exists. Again, this applies to every Q position at which the low-energy magnetic signal is seen. It may also be necessary for the theoretical calculations based on the Heisenberg model on the kagome lattice to be modified to match more closely the spin Hamiltonian of herbertsmithite.

The observed spinon continuum is the strongest evidence yet that the ground state of the $S = 1/2$ kagome antiferromagnet herbertsmithite is a quantum spin liquid. The measured spin correlations are short ranged; however, no spin gap is observed. An intriguing

aspect of quantum spin liquids is that whereas the spin correlations may be short ranged, the quantum coherence is long ranged. These neutron results serve as a strong foundation for detailed tests of theoretical proposals for spin liquid states on the kagome lattice.

METHODS SUMMARY

Single-crystal samples of $\text{ZnCu}_3(\text{OD})_6\text{Cl}_2$ were grown hydrothermally in tube furnaces under a temperature gradient¹⁷. Fifteen of the largest crystals were co-aligned on an aluminium sample holder, yielding a total mass of 1.2 g. The overall mosaic (the full-width at half-maximum of the angular distribution of crystallites that comprise the sample) was determined by neutron diffraction to be $\sim 2^\circ$. An identical aluminium sample holder was prepared for the purpose of background subtraction. Inelastic neutron scattering experiments were performed using the multi-axis crystal spectrometer at the NIST Center for Neutron Research. The incident neutron energy was selected using a doubly focused pyrolytic graphite monochromator. The final analysed neutron energy was fixed to be either $E_f = 5.1$ or 3.0 meV , for energy resolutions of 0.21 meV (half-width at half-maximum) and 0.08 meV , respectively. The sample was mounted in the (H, K, 0) scattering plane, and a pumped helium cryostat with a base temperature of $T = 1.6 \text{ K}$ was used to control the temperature of the sample.

Received 3 June; accepted 2 October 2012.

- Balents, L. Spin liquids in frustrated magnets. *Nature* **464**, 199–208 (2010).
- Anderson, P. W. The resonating valence bond state in La_2CuO_4 and superconductivity. *Science* **235**, 1196–1198 (1987).
- Lee, P. A., Nagaosa, N., & Wen, X.-G. Doping a Mott insulator: physics of high-temperature superconductivity. *Rev. Mod. Phys.* **78**, 17–85 (2006).
- Ioffe, L. B. *et al.* Topologically protected quantum bits using Josephson junction arrays. *Nature* **415**, 503–506 (2002).
- Tennant, D. A., Cowley, R. A., Nagler, S. E. & Tsvetlik, A. M. Measurement of the spin excitation continuum in one-dimensional KCuF_3 using neutron scattering. *Phys. Rev. B* **52**, 13368–13380 (1995).
- Sachdev, S. Kagome- and triangular-lattice Heisenberg antiferromagnets: ordering from quantum fluctuations and quantum-disordered ground states with unconfined bosonic spinons. *Phys. Rev. B* **45**, 12377–12396 (1992).
- Elser, V. Nuclear antiferromagnetism in a registered ^3He solid. *Phys. Rev. Lett.* **62**, 2405–2408 (1989).
- Lecheminant, P., Bernu, B., Lhuillier, C., Pierre, L. & Sindzingre, P. Order versus disorder in the quantum Heisenberg antiferromagnet on the kagome lattice using exact spectra analysis. *Phys. Rev. B* **56**, 2521–2529 (1997).
- Shores, M. P., Nytko, E. A., Bartlett, B. M. & Nocera, D. G. A structurally perfect $S = 1/2$ kagome antiferromagnet. *J. Am. Chem. Soc.* **127**, 13462–13463 (2005).

10. Freedman, D. E. *et al.* Site specific x-ray anomalous dispersion of the geometrically frustrated kagomé magnet herbertsmithite, $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$. *J. Am. Chem. Soc.* **132**, 16185–16190 (2010).
11. Helton, J. S. *et al.* Spin dynamics of the spin-1/2 kagome lattice antiferromagnet $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$. *Phys. Rev. Lett.* **98**, 107204 (2007).
12. Mendels, P. *et al.* Quantum magnetism in the paratacamite family: towards an ideal kagomé lattice. *Phys. Rev. Lett.* **98**, 077204 (2007).
13. de Vries, M. A. *et al.* Scale-free antiferromagnetic fluctuations in the $S = 1/2$ kagome antiferromagnet herbertsmithite. *Phys. Rev. Lett.* **103**, 237201 (2009).
14. Zorko, A. *et al.* Dzyaloshinsky-Moriya anisotropy in the spin-1/2 kagome compound $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$. *Phys. Rev. Lett.* **101**, 026405 (2008).
15. Han, T., Chu, S. & Lee, Y. S. Refining the spin Hamiltonian in the spin-1/2 kagome lattice antiferromagnet $\text{ZnCu}_3(\text{OH})_6\text{Cl}_2$ using single crystals. *Phys. Rev. Lett.* **108**, 157202 (2012).
16. Yan, S., Huse, D. A. & White, S. R. Spin liquid ground state of the $S = 1/2$ kagome Heisenberg antiferromagnet. *Science* **332**, 1173–1176 (2011).
17. Han, T. H. *et al.* Synthesis and characterization of single crystals of the spin-1/2 kagome-lattice antiferromagnet $\text{Zn}_{1-x}\text{Cu}_x(\text{OH})_6\text{Cl}_2$. *Phys. Rev. B* **83**, 100402(R) (2011).
18. Imai, T., Fu, M., Han, T. H. & Lee, Y. S. Local spin susceptibility of the $S = 1/2$ kagome lattice in $\text{ZnCu}_3(\text{OD})_6\text{Cl}_2$. *Phys. Rev. B* **84**, 020411(R) (2011).
19. Ofer, O., Keren, A., Brewer, J. H., Han, T. H. & Lee, Y. S. The herbertsmithite Hamiltonian: μSR measurements on single crystals. *J. Phys. Condens. Matter* **23**, 164207 (2011).
20. Wulferding, D. *et al.* Interplay of thermal and quantum spin fluctuations in the kagome lattice compound herbertsmithite. *Phys. Rev. B* **82**, 144412 (2010).
21. Birgeneau, R. J. *et al.* Instantaneous spin correlations in La_2CuO_4 . *Phys. Rev. B* **59**, 13788–13794 (1999).
22. Grohol, D. *et al.* Spin chirality on a two-dimensional frustrated lattice. *Nature Mater.* **4**, 323–328 (2005).
23. Kivelson, S. A., Rokhsar, D. S. & Sethna, J. P. Topology of the resonating valence-bond state: solitons and high- T_c superconductivity. *Phys. Rev. B* **35**, 8865–8868 (1987).
24. Lovesey, S. W. *Theory of Neutron Scattering from Condensed Matter* (Clarendon, 1984).
25. Helton, J. S. *et al.* Dynamic scaling in the susceptibility of the spin-1/2 Kagome lattice antiferromagnet herbertsmithite. *Phys. Rev. Lett.* **104**, 147201 (2010).
26. Matan, K. *et al.* Spin waves in the frustrated kagomé lattice antiferromagnet $\text{KFe}_3(\text{OH})_6(\text{SO}_4)_2$. *Phys. Rev. Lett.* **96**, 247201 (2006).
27. Hao, Z. & Tchernyshyov, O. Structure factor of low-energy spin excitations in a $S = 1/2$ kagome antiferromagnet. *Phys. Rev. B* **81**, 214445 (2010).
28. Singh, R. R. P. & Huse, D. A. Triplet and singlet excitations in the valence bond crystal phase of the kagome lattice Heisenberg model. *Phys. Rev. B* **77**, 144415 (2008).
29. Ran, Y., Hermele, M., Lee, P. A. & Wen, X.-G. Projected wavefunction study of spin-1/2 Heisenberg model on the kagome lattice. *Phys. Rev. Lett.* **98**, 117205 (2007).
30. Ryu, S., Motrunich, O. I., Alicea, J. & Fisher, M. P. A. Algebraic vortex liquid theory of a quantum antiferromagnet on the kagome lattice. *Phys. Rev. B* **75**, 184406 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge T. Senthil, P. A. Lee, Z. Hao and O. Tchernyshyov for discussions, and J. Wen for assistance with data reduction. The work at MIT was supported by the US Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences under grant no. DE-FG02-07ER46134. This work used facilities supported in part by the US National Science Foundation under agreement no. DMR-0944772. The work at IQM was supported by the DOE, Office of Basic Energy Sciences, Division of Material Sciences and Engineering under award DE-FG02-08ER46544.

Author Contributions Y.S.L. supervised all aspects of the research. T.-H.H. synthesized and characterized the sample. T.-H.H., J.A.R.-R., J.S.H. and C.B. collected the neutron scattering data. T.-H.H. analysed the data with input from Y.S.L., C.B., J.S.H. and J.A.R.-R. S.C. and D.G.N. aided in the sample preparation process. T.-H.H. and Y.S.L. wrote the manuscript with comments from all others. The manuscript reflects the contributions of all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.S.L. (younglee@mit.edu) or T.-H.H. (tianheng@alum.mit.edu)

Quasi-cylindrical wave contribution in experiments on extraordinary optical transmission

Frerik van Beijnum¹, Chris R  tief², Chris B. Smiet¹, Haitao Liu³, Philippe Lalanne⁴ & Martin P. van Exter¹

A metal film perforated by a regular array of subwavelength holes shows unexpectedly large transmission at particular wavelengths, a phenomenon known as the extraordinary optical transmission (EOT) of metal hole arrays¹. EOT was first attributed to surface plasmon polaritons, stimulating a renewed interest in plasmonics^{2–4} and metallic surfaces with subwavelength features^{5–7}. Experiments soon revealed that the field diffracted at a hole or slit is not a surface plasmon polariton mode alone⁸. Further theoretical analysis⁹ predicted that the extra contribution, from quasi-cylindrical waves^{10–13}, also affects EOT. Here we report the experimental demonstration of the relative importance of surface plasmon polaritons and quasi-cylindrical waves in EOT by considering hole arrays of different hole densities. From the measured transmission spectra, we determine microscopic scattering parameters which allow us to show that quasi-cylindrical waves affect EOT only for high densities, when the hole spacing is roughly one wavelength. Apart from providing a deeper understanding of EOT, the determination of microscopic scattering parameters from the measurement of macroscopic optical properties paves the way to novel design strategies.

Understanding EOT quantitatively has been one of the main challenges in plasmonics in the past decade (see ref. 7 for a review). An important contribution to EOT may be the quasi-cylindrical wave (QCW), which is the field diffracted at a subwavelength indentation on the metal surface, simultaneously with the surface plasmon polariton (SPP) contribution^{10,11,13}. The justification for separating these two field contributions lies in the fact that QCWs exist in the absence of SPPs, that is, for a perfect electric conductor interface^{10,12,14} and for a dielectric interface¹⁵. For a perfect electric conductor, the QCW is just a cylindrical wave in free space with a $1/x^{1/2}$ decay rate, x being the distance to the scatterer. For metals of finite conductivity at optical frequencies, the cylindrical-wave behaviour is seen only in the vicinity of the indentation, within a distance of a few wavelengths. Farther from the indentation, the QCW decay rate increases until it reaches an asymptotic algebraic value of $1/x^{3/2}$ at large distances ($x \rightarrow \infty$). Because of the periodicity of hole arrays, all these decay rates are simultaneously present in EOT.

Another important result in our understanding of EOT was the development of a semi-analytic SPP model, which quantifies the SPP contribution to EOT. This microscopic model relies on scattering parameters that describe the interaction between the incident light and the holes. The microscopic model correctly predicts the EOT line shape but underestimates the magnitude of the transmission by a factor of roughly two for visible frequencies. These results suggest that SPPs are responsible for only about 50% of EOT⁹, with the other 50% due to QCWs. So far, this interpretation is purely theoretical and remains conceptual, and we are not aware of any experimental confirmation. The likely reason for this is that many scattering parameters in the model are dispersive and had to be calculated using complicated simulations⁹. Measuring all these parameters with sufficient accuracy is a tremendous experimental challenge that we address here.

In this Letter, we provide direct evidence for the respective roles of SPPs and QCWs in EOT, by exploiting the fact that they have different characteristic damping lengths. We do this by measuring the transmission spectra of a series of hole arrays with varying hole densities, designed to resonate at the same near-infrared frequency (~ 750 nm). We find that the normalized transmission peaks of all the low-density arrays are almost identical. The array of highest density, which corresponds to the classical hole array in ref. 1, has a hole spacing that corresponds to a QCW damping length of one wavelength. This array shows a pronounced increase in normalized transmission compared with the transmission of the other arrays. The observation of this atypical behaviour for the highest-density array is direct proof of the involvement of QCWs in EOT, or more generally of the QCWs in phenomena related to subwavelength metallic gratings.

Figure 1a illustrates how we vary the hole density: we increase the distance between ‘hole chains’ (lines of holes whose separation in the y direction, $a_y = a_0 = 450$ nm, we do not vary) while keeping the properties of these chains fixed. The distance between the chains is chosen to be $a_x = qa_0$, where $q = 1, 2, 3, 4, 6$ and 7 . The holes perforate two metal layers: a 150-nm gold layer that is deposited on a glass substrate, and a 20-nm chrome layer that damps the SPP at the air interface (Fig. 1b). Figure 1c shows a scanning electron microscope image of the $q = 2$ hole array. We measure the zeroth-order transmission spectra of each array using light polarized in the x direction. As a result, SPPs propagating along the y axis are negligible^{9,16–18}.

Figure 2a shows the measured spectra on a semi-log scale. Each array has a transmission peak between 735 and 775 nm, which shifts to shorter wavelengths as the hole spacing increases. The transmission minima of these resonances are at 724 ± 4 nm, which is very close to 729 nm, the expected value of $\text{Re}(n_{\text{eff}})a_0$, where n_{eff} is the ratio between the magnitude of the SPP wavevector and that of the free-space wavevector^{19,20}. The magnitudes of the transmission minima range between 0.6×10^{-4} and 2×10^{-4} . The resonance peaks are the result of a surface wave propagating in the x direction, with the number of oscillations per lattice period equal to the value of q for the hole array. For $q = 4, 6$ and 7 , additional resonances are also visible at wavelengths between 800 and 1,000 nm. These resonances correspond to a surface wave propagating in the x direction with $q - 1$ oscillations per period.

As evidenced by the log-scale plots, the zeroth-order transmission drastically decreases as the hole density decreases. As will be explained below (equation (2)), it essentially scales as $1/q^2$. Thus, in Fig. 2b we plot the product of q^2 and transmission on a linear scale, for $q = 1, 2, 4$ and 6 . The results for $q = 3$ and 7 have been removed for clarity. By plotting the data on a linear scale, we see that the scaled transmission is almost constant when going from $q = 2$ to $q = 6$. However, for $q = 1$ the transmission peak is markedly larger, being more than two times larger than all the other peaks. The fact that all the arrays except $q = 1$ have almost identical extraordinary transmission peaks is not coincidental and, as will be explained below, reveals the resonant transmission due to SPPs only. In this respect, the sudden and marked

¹Leiden University, Huygens Laboratory, PO Box 9504, 2300 RA Leiden, The Netherlands. ²FOM Institute for Atomic and Molecular Physics, Science Park 104, 1098 XG Amsterdam, The Netherlands. ³Key Laboratory of Optical Information Science and Technology, Ministry of Education, Institute of Modern Optics, Nankai University, Tianjin 300071, China. ⁴Laboratoire Photonique, Num  rique et Nanosciences (LP2N), Institut d’Optique, Universit   Bordeaux 1, CNRS, 33405 Talence Cedex, France.

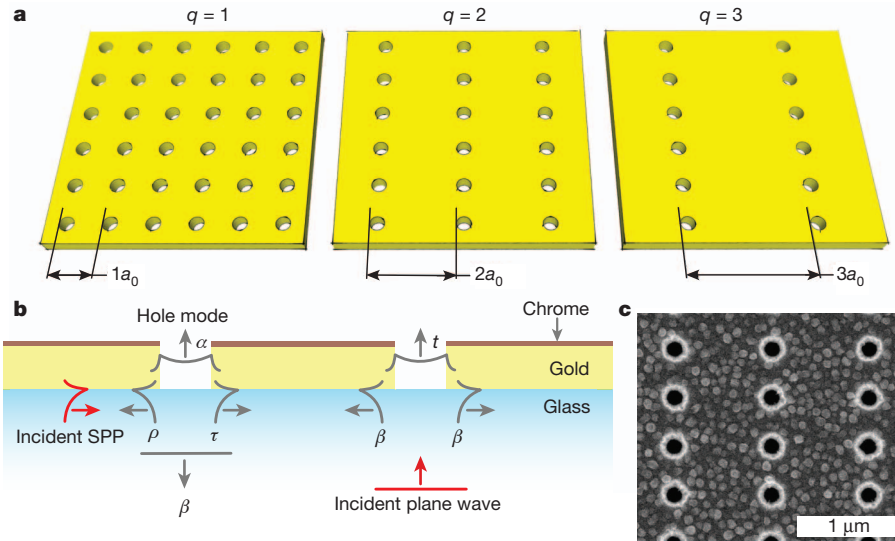


Figure 1 | Sample design. To demonstrate a QCW contribution to EOT, we compare the transmissions of a series of hole arrays that have different densities and all resonate at the same wavelength. **a**, To decrease the hole density, the distance between hole chains is increased. **b**, The relevant scattering processes. At the left hole chain, there is an incident SPP (red arrow) that is transmitted

increase of the peak transmission for the highest density ($q = 1$) is one of our main results: it is a direct signature of the additional short-range contribution of the QCWs to EOT.

The previous conclusion is inferred qualitatively on the basis of the difference in damping lengths between SPPs and QCWs. Using the microscopic SPP model in ref. 9, we now aim at providing a fully quantitative analysis of the data in Fig. 2 to support our conclusions explicitly. The main drawback of the microscopic approach is that it requires knowledge of many dispersive, complex-valued parameters, which need to be calculated using full-vectorial simulations at every wavelength. This precludes its direct application to the analysis of experimental data. Below we discuss the model in depth and show that an analytical expression involving only five independent, real, non-dispersive, fitted parameters accurately reproduces the whole set of experimental spectra in Fig. 2 for $q \geq 2$, but produces wrong predictions for $q = 1$.

The microscopic model, as originally formulated in ref. 9, considers one-dimensional hole chains (along the y direction) to be the elementary scatterers of the hole arrays and assumes that the electromagnetic interaction between the hole chains is mediated only by the SPP modes on unperturbed interfaces, the QCWs being totally neglected. The microscopic model couples three modes: the SPP mode at the gold-glass interface, which consists of SPPs propagating in both directions along the x axis; the fundamental mode within the holes; and the free-space modes.

The modes are coupled by the following scattering parameters (Fig. 1b). An incident SPP can couple into the holes of the hole chain (α), couple to free space (β), or be transmitted (τ) or reflected (ρ). An incident plane wave can couple to a SPP (β) or directly into the holes of the hole chain (t). The propagation of the SPP from one hole chain to another is described by $e^{-ik_{\text{spp}}a}$, where $k_{\text{spp}} = 2\pi n_{\text{eff}}/\lambda$ is the wavevector magnitude of the SPP on a flat gold-glass interface and a is the distance between the hole chains. Because we keep the hole size constant for each array, the scattering parameters (α , β , ρ , τ and t) should be the same for each array. Only the hole spacing a_x depends on the parameter q : $a_x = qa_0$.

In line with the coupled-mode approach proposed in ref. 9, we derive (Supplementary Information) the transmission coefficient, t_t , of a single hole chain, modified by the SPPs excited at the surrounding hole chains:

(τ), reflected (ρ), coupled into the hole (α) or coupled to free space (β). At the right hole chain, there is an incident free-space mode (red arrow) that scatters into a SPP (β) or couples into the hole (t). **c**, Scanning electron microscope picture of the $q = 2$ hole array.

$$t_t = t + \frac{2\alpha\beta}{e^{-ik_{\text{spp}}a} - (\rho + \tau)} \quad (1)$$

The assumption that the hole chains excite only a plane SPP wave in the x direction is valid if the incident polarization is along the x axis. Moreover, the holes within the hole chain have to be less than a SPP wavelength apart⁹. The model neglects the influence of the metal-air interface. In Supplementary Information, we show that this assumption holds if the SPPs at that interface are sufficiently damped by the chrome layer.

The optical transmission measured at a particular angle is a sum of the field contributions from all illuminated hole chains. If we define N_0 as the number of illuminated hole chains for the $q = 1$ array, $N = N_0/q$ is the number of hole chains as a function of q . In the zeroth diffraction order, the contributions from all hole chains are in phase, and the transmitted intensity is hence²¹

$$T \propto 1/q^2 |t_t|^2 \quad (2)$$

where the proportionality factor is independent of q (Supplementary Information).

Equations (1) and (2) are not yet simple enough to be used to model experimental data. The six scattering parameters are complex, such that there are twelve real-valued parameters. To reduce the number of free parameters, we combine α and β into $\alpha\beta$ and ρ and τ into $\rho + \tau$. Furthermore, we need only consider the phase difference between t and $\alpha\beta$. For the complex variable k_{spp} , we combine previously measured, linearly interpolated values for the gold properties²² with a measured value of the index of refraction of the substrate (1.51). This leaves us with a model with five free parameters.

Another crucial step towards fitting the model to the data is the wavelength dependence of the parameters. We use the following approximations: $t \propto \lambda^{-2}$, as indicated by several studies^{7,23–25}; $\alpha\beta \propto \lambda^{-4}$, in rough agreement with fourfold reduction in the magnitude of $\alpha\beta$ for a wavelength increase of 38% (ref. 9); we keep $\rho + \tau$ constant, as calculations show that this sum has limited dispersion; the phase of $\rho + \tau$ and the phase difference between $\alpha\beta$ and t are also kept constant as functions of wavelength. As a result, we have the following fit equation:

$$T = q^{-2} \left| \frac{\lambda_0^2}{\lambda^2} p_1 + \frac{\lambda_0^4}{\lambda^4} \frac{p_2 e^{ip_3}}{e^{-ik_{\text{spp}}qa_0} - p_4 e^{ip_5}} \right|^2 \quad (3)$$

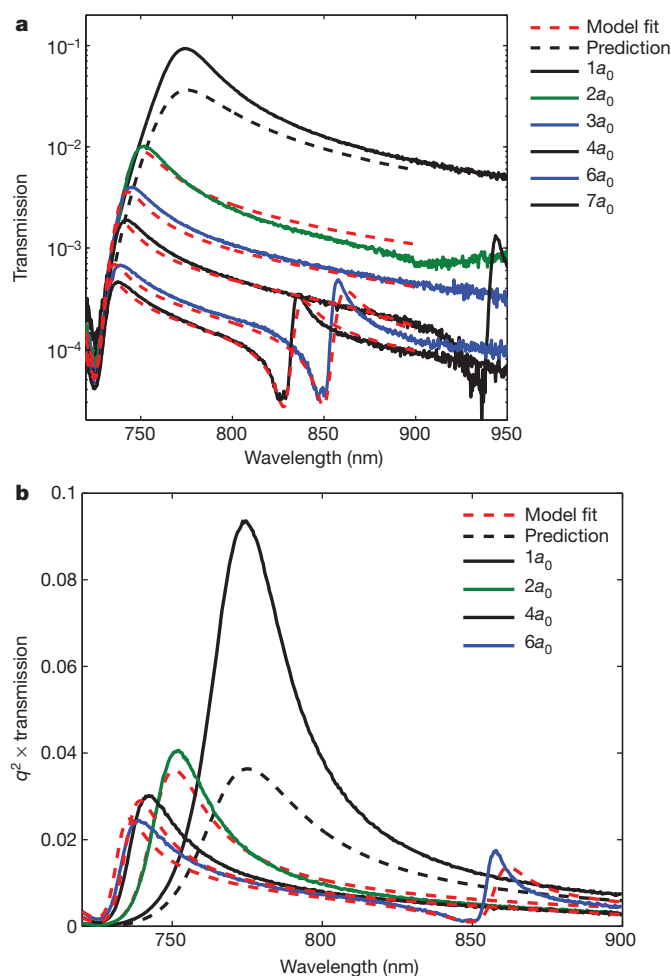


Figure 2 | Measured transmission spectra (solid curves) and the fitted SPP model (red dashed curves). Each dashed curve has the same parameters except for the q value. The parameters are based on a fit of equation (3) to the data of arrays with $q \geq 2$. The black dashed curve is a model prediction, plotting equation (3) with the same parameters as the other curves but with $q = 1$. **a**, The model accurately predicts all maxima and minima of the measured data for $q \geq 2$, showing that the transmission of these arrays is dominated by SPPs. **b**, Plotting the product of q^2 and transmission shows that there is a marked increase in transmission for $q = 1$ relative to the other arrays. By also plotting the SPP model, we show that this behaviour is not predicted by considering SPPs alone.

where $\lambda_0 = 800$ nm and p_1 to p_5 are the five real-valued parameters in the model. The physical significance of parameters p_1 and p_4 is discussed in Supplementary Information.

The fitted curves obtained with the model are shown with the red dashed curves in Fig. 2. For the fit, we only used the data of arrays $q = 2$ to $q = 7$, as we expect the transmission of these arrays to be dominated by SPPs. To ensure that the minima are fitted well, we fitted the base-10 log of the model to that of the data. The result is plotted in Fig. 2a, and the fit values are given in Table 1. For the arrays with $q = 2, 3, 4, 6$ and 7 , the model describes the spectral positions of all minima and maxima very accurately, including the $(q - 1, 0)$ resonances that appear close to the wavelength of 850 nm. The transmission magnitude also is well modelled.

Notably, equation (3) (Fig. 2, dashed black curves) fails to predict the experimental data for $q = 1$. In particular, it underestimates the transmission peak by a factor of 2.5. The latter is consistent with the theoretical predictions in ref. 9, which conclude that SPPs account for only half of the total transmitted energy at peak transmittance for a resonance wavelength of 700 nm. In Supplementary Information, we expand the SPP microscopic model to incorporate QCWs (see

Table 1 | The values of the fitting parameters p_1 to p_5

Parameter	Fit value	Error
p_1	0.0740	0.0006
p_2	0.0243	0.0006
p_3	0.681π	0.013π
p_4	0.857	0.009
p_5	0.132π	0.006π
$ \exp(-ik_{\text{spp}}a_0) $	1.021	NA

For completeness, we also give the value of $|\exp(-ik_{\text{spp}}a_0)|$, evaluated at 800 nm. We also give an estimate of the error in the fit values. NA, not applicable.

equation (4) therein), and this extended model accurately describes the experimental data for $q = 1$ (Supplementary Fig. 3). We emphasize that this result is achieved using the same scattering parameters and does not require any additional fitting.

The observation that the transmission spectra of these six different hole arrays (including the particular $q = 1$ case) can be described with only five parameters is a considerable success of the microscopic theory. Moreover, it is of importance that these parameters are directly related to elementary scattering processes. It seems that the resonance of a perforated metal surface can be modelled with a combination of an SPP mode and a QCW on an unperturbed surface, taking into account a few extra scattering parameters, which can be inferred from simple transmission measurements.

Our experiments quantitatively show the respective contributions of SPPs and QCWs to EOT. The main step in our analysis was the derivation of the elementary scattering parameters defined in ref. 9 from classical transmission measurements performed on a series of hole arrays. The possibility of making these microscopic scattering parameters experimentally accessible may be paramount in understanding and designing complex periodic or aperiodic metallic structures. Hence, variations on this modelling approach may have important implications in applying plasmonic structures to sensors, photovoltaics, light-emitting diodes and lasers.

METHODS SUMMARY

The sample was fabricated as follows. Using electron-beam lithography, we created pillars on a glass substrate. Then we deposited 150 nm of gold and, subsequently, 20 nm of chrome. Finally, we etched away the pillars. Using scanning electron microscope pictures of each array, we measured the average hole radius to be 81 ± 4 nm. By studying the position of the diffraction orders using a 635-nm laser diode, we found that for all arrays a_0 has the same value to within 2 nm. Each array covers an area of $400 \mu\text{m} \times 400 \mu\text{m}$.

The transmission spectra were measured using the following set-up. The light of a halogen lamp was filtered (long pass, 600 nm) and coupled to a 200- μm -diameter fibre. The light coupled out of the fibre was polarized along the x axis. The end facet of the fibre was imaged onto the sample with a magnification of 1.5. The 300- μm -diameter spot was centred on the $400 \mu\text{m} \times 400 \mu\text{m}$ arrays. The zeroth-order transmission was imaged (magnification, 2/3) onto a 365- μm -diameter fibre that led to an Ocean Optics 2000+ USB spectrometer. To measure the reference spectrum, we moved the sample out of the beam using a stage.

An accurate measurement of the transmission spectrum requires sufficient spatial and temporal coherence of the illuminating and detected light. The spatial coherence at the sample was guaranteed by the limited numerical aperture ($\sim (6 \pm 2) \times 10^{-3}$) of the illumination. This corresponded to a coherence length of a few tens of micrometres on the sample²¹, which is much longer than an SPP propagation length. Temporal coherence was ensured by the spectral resolution of the detecting spectrometer. The 1-nm resolution corresponded to coherence times of hundreds of femtoseconds²¹ and was an order of magnitude finer than strictly needed.

Received 6 June; accepted 4 October 2012.

1. Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T. & Wolff, P. A. Extraordinary optical transmission through sub-wavelength hole arrays. *Nature* **391**, 667–669 (1998).
2. Yanik, A. A. et al. An optofluidic nanoplasmonic biosensor for direct detection of live viruses from biological media. *Nano Lett.* **10**, 4962–4969 (2010).
3. Oulton, R. F. et al. Plasmon lasers at deep subwavelength scale. *Nature* **461**, 629–632 (2009).
4. Atwater, H. A. & Polman, A. Plasmonics for improved photovoltaic devices. *Nature Mater.* **9**, 205–213 (2010).

5. Barnes, W. L., Dereux, A. & Ebbesen, T. W. Surface plasmon subwavelength optics. *Nature* **424**, 824–830 (2003).
6. García de Abajo, F. J. Light scattering by particle and hole arrays. *Rev. Mod. Phys.* **79**, 1267–1290 (2007).
7. García-Vidal, F. J., Martín-Moreno, L., Ebbesen, T. W. & Kuipers, L. Light passing through subwavelength apertures. *Rev. Mod. Phys.* **82**, 729–787 (2010).
8. Gay, G. *et al.* The optical response of nanostructured surfaces and the composite diffracted evanescent wave model. *Nature Phys.* **2**, 262–267 (2006).
9. Liu, H. T. & Lalanne, P. Microscopic theory of the extraordinary optical transmission. *Nature* **452**, 728–731 (2008).
10. Lalanne, P. & Hugonin, J. P. Interaction between optical nano-objects at metallo-dielectric interfaces. *Nature Phys.* **2**, 551–556 (2006).
11. Dai, W. & Soukoulis, C. M. Theoretical analysis of the surface wave along a metal-dielectric interface. *Phys. Rev. B* **80**, 155407 (2009).
12. Lalanne, P., Hugonin, J. P., Liu, H. T. & Wang, B. A microscopic view of the electromagnetic properties of sub- λ metallic surfaces. *Surf. Sci. Rep.* **64**, 453–469 (2009).
13. Nikitin, A. Y., García-Vidal, F. J. & Martín-Moreno, L. Surface electromagnetic field radiated by a subwavelength hole in a metal film. *Phys. Rev. Lett.* **105**, 073902 (2010).
14. Nikitin, A., Yu, Rodrigo, S. G., García-Vidal, F. J. & Martín-Moreno, L. In the diffraction shadow: Norton waves versus surface plasmon polaritons in the optical region. *N. J. Phys.* **11**, 123020 (2009).
15. Gan, C. H., Lalouat, L., Lalanne, P. & Aigouy, L. Optical quasicylindrical waves at dielectric interfaces. *Phys. Rev. B* **83**, 085422 (2011).
16. Barnes, W. L., Murray, W. A., Dintinger, J., Devaux, E. & Ebbesen, T. W. Surface plasmon polaritons and their role in the enhanced transmission of light through periodic arrays of subwavelength holes in a metal film. *Phys. Rev. Lett.* **92**, 107401 (2004).
17. Stolkwijk, D. *et al.* Enhanced coupling of plasmons in hole arrays with periodic dielectric antennas. *Opt. Lett.* **33**, 363–365 (2008).
18. Yin, L. *et al.* Subwavelength focusing and guiding of surface plasmons. *Nano Lett.* **5**, 1399–1402 (2005).
19. Lalanne, P., Sauvan, C., Hugonin, J. P., Rodier, J. C. & Chavel, P. Perturbative approach for surface plasmon effects on flat interfaces periodically corrugated by subwavelength apertures. *Phys. Rev. B* **68**, 125404 (2003).
20. Maystre, D., Fehrembach, A.-L. & Popov, E. Plasmonic antiresonance through subwavelength hole arrays. *J. Opt. Soc. Am. A* **28**, 342–355 (2011).
21. Born, M. & Wolf, E. *Principles of Optics* (Pergamon, 1986).
22. Palik, E. D. *Handbook of Optical Constants of Solids* 286–295 (Academic, 1985).
23. Bethe, H. A. Theory of diffraction by small holes. *Phys. Rev.* **66**, 163–182 (1944).
24. van der Molen, K. L. *et al.* Role of shape and localized resonances in extraordinary transmission through periodic arrays of subwavelength holes: experiment and theory. *Phys. Rev. B* **72**, 045421 (2005).
25. van Beijnum, F., Rétif, C., Smiet, C. B. & van Exter, M. P. Transmission processes in random patterns of subwavelength holes. *Opt. Lett.* **36**, 3666–3668 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge E. R. Eliel for discussions. H.L. acknowledges a Poste Rouge fellowship from CNRS and the 973 Program (2013CB328701). This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO).

Author Contributions F.v.B. was the primary researcher on the project; he designed the experiment and wrote the paper. With C.B.S., F.v.B. conducted the experiments and analysed the data. C.R. made the samples. H.L., P.L. and M.P.v.E. made essential contributions to interpreting the results and writing the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.v.B. (beijnum@physics.leidenuniv.nl).

Laboratory measurements of the viscous anisotropy of olivine aggregates

L. N. Hansen^{1†}, M. E. Zimmerman¹ & D. L. Kohlstedt¹

A marked anisotropy in viscosity develops in Earth's mantle as deformation strongly aligns the crystallographic axes of the individual grains that comprise the rocks. On the basis of geodynamic simulations, processes significantly affected by viscous anisotropy include post-glacial rebound^{1,2}, foundering of lithosphere³ and melt production above subduction zones⁴. However, an estimate of the magnitude of viscous anisotropy based on the results of deformation experiments on single crystals⁵ differs by three orders of magnitude from that obtained by grain-scale numerical models of deforming aggregates with strong crystallographic alignment^{6–8}. Complicating matters, recent experiments indicate that deformation of the uppermost mantle is dominated by dislocation-accommodated grain-boundary sliding⁹, a mechanism not activated in experiments on single crystals and not included in numerical models. Here, using direct measurements of the viscous anisotropy of highly deformed polycrystalline olivine, we demonstrate a significant directional dependence of viscosity. Specifically, shear viscosities measured in high-strain torsion experiments are 15 times smaller than normal viscosities measured in subsequent tension tests performed parallel to the torsion axis. This anisotropy is approximately an order of magnitude larger than that predicted by grain-scale simulations. These results indicate that dislocation-accommodated grain-boundary sliding produces an appreciable anisotropy in rock viscosity. We propose that crystallographic alignment imparts viscous anisotropy because the rate of deformation is limited by the movement of dislocations through the interiors of the crystallographically aligned grains. The maximum degree of anisotropy is reached at geologically low shear strain (of about ten) such that deforming regions of the upper mantle will exhibit significant viscous anisotropy.

Many geodynamic processes are directly affected by the development of viscous anisotropy (VA) in Earth's mantle. Numerical simulations indicate that appreciable VA in mantle rocks greatly modifies melt production from mantle wedges above subduction zones⁴, the spatial and temporal frequency of lithospheric density instabilities³, the aspect ratio and boundary-layer thickness of mantle convection cells^{1,10} and the dynamics of post-glacial rebound^{1,2}. Furthermore, the correlation between the orientation of continental rifts and pre-existing mantle fabrics suggests that VA plays a key role in plate-boundary formation¹¹.

The incorporation of VA in numerical simulations is often simplified by reducing the problem to two dimensions and assuming the viscosity is Newtonian and the material is transversely isotropic. Deformation can then be described by^{1,10}

$$\begin{bmatrix} \sigma_n \\ \sigma_s \end{bmatrix} = \begin{bmatrix} \eta_n & 0 \\ 0 & \eta_s \end{bmatrix} \begin{bmatrix} \dot{\epsilon}_n \\ \dot{\epsilon}_s \end{bmatrix} \quad (1)$$

where σ is deviatoric stress, η is viscosity, $\dot{\epsilon}$ is strain rate, and the subscripts n and s refer to axial deformation normal to the plane of isotropy and shear deformation parallel to the plane of isotropy, respectively (further detail is given in Supplementary Information).

The magnitude of anisotropy, δ , is then defined as $\delta = \eta_n/\eta_s$ such that $\delta = 1$ if the material is isotropic. Geodynamic simulations have employed values of δ ranging from 1.5 to 100 (refs 1, 3, 4).

However, despite the appreciable influence of VA on geodynamic processes, the magnitude of anisotropy in mantle rocks is uncertain. One of the most important sources of VA in plastically deforming rocks is the formation of strong crystallographic preferred orientations (CPOs). Most estimates of VA in an olivine aggregate with a strong CPO are based on creep data from oriented single crystals of San Carlos olivine, which indicate that a maximum value of $\delta > 1,000$ is achieved if the viscosity of crystals oriented for axial compression normal to the (010) plane is compared to that of crystals oriented for shear on the weakest slip system, (010)[100] (ref. 5). Several researchers have explored the VA of olivine aggregates arising from CPO development by incorporating results from single-crystal experiments in grain-scale numerical models. Although extrapolating rheological behaviour from single crystals to polycrystalline aggregates is a complicated mechanical problem, predictions of the development of CPOs and of the evolution of strength in olivine have been made using various types of self-consistent estimates^{6–8,12,13}. Notably, multiscale modelling of olivine aggregates with strong initial CPOs^{7,8} indicates maximum anisotropies of $\delta = 1.5$ –3.0. However, these simulations do not account for deformation by sliding along grain boundaries or climb of dislocations; instead, they resort to the use of fictitious slip systems to maintain strain compatibility. This simplification is clearly problematic in light of recent laboratory-based investigations of olivine deformation^{9,14}, which demonstrate that dislocation-accommodated grain-boundary sliding in olivine is an important (if not the dominant) deformation mechanism in the upper mantle. Additionally, to promote computational stability, multiscale models assessing the VA of olivine aggregates have been limited to relatively weak crystallographic fabrics.

Motivated by the large discrepancy in estimates of δ for deformed olivine-rich rocks, we utilize an innovative laboratory-based approach to directly measure the magnitude and evolution of shear and normal viscosities in olivine aggregates with a developing CPO. Thin-walled cylinders of the sample material were initially deformed in torsion to induce microstructural anisotropy (Fig. 1). As observed in prior high-strain experiments on olivine^{15,16}, the average orientations of the (010) planes and [100] axes roughly align with the shear plane and shear direction, respectively, as deformation progresses, indicating that intragranular deformation was largely accommodated by slip on the (010)[100] slip system. In deformed thin-walled cylinders, the microstructure (CPO and grain size) is essentially homogeneous throughout the sample because the radial gradients in stress and strain are relatively small¹⁷. Samples were subsequently deformed in tension such that the axial load was parallel to both the torsion axis and the dominant orientation of the [010] axes (further details are given in Methods). Thus, the viscosities measured in torsion and tension reflect the minimum and maximum values of η_s and η_n , respectively, based on a comparison of (1) the viscosity of olivine single crystals oriented to

¹Department of Earth Sciences, University of Minnesota, 310 Pillsbury Drive SE, Minneapolis, Minnesota 55455, USA. †Present address: Department of Geological and Environmental Sciences, Stanford University, 450 Serra Mall, Bldg 320, Stanford, California 94305, USA.

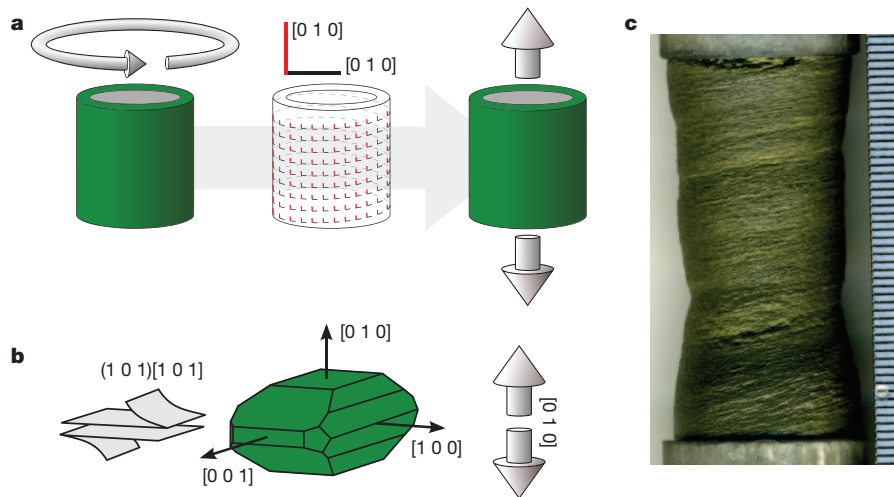


Figure 1 | Overview of experiment design. **a**, Schematic diagrams of thin-walled cylinders of olivine deformed in torsion (left) and in tension (right). Green indicates the olivine sample and grey indicates the central Ni cylinder. The central Ni cylinder is weak relative to the olivine aggregate. The centre diagram illustrates the dominant crystallographic orientation as a function of

position around the cylinder. Crystallographic axes are indicated. **b**, Schematic diagrams of the orientation of the applied stress with respect to the dominant crystal orientation. **c**, Photograph of sample PT0652 after torsion and tension tests. The Ni and Fe jackets have been removed. Tick marks are 0.5 mm.

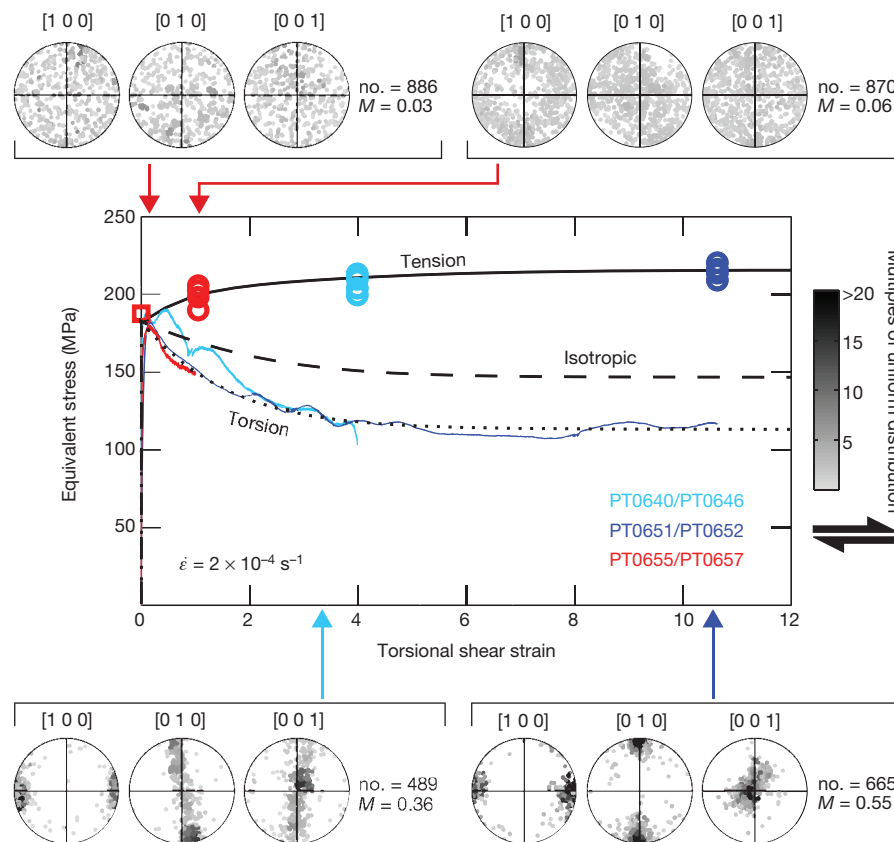


Figure 2 | Mechanical and crystallographic data from entire data set. Main panel, coloured curves are data from torsion experiments, and coloured data points are data from tension experiments normalized to the same effective strain rate as used in torsion experiments (see Supplementary Information). Data from torsion and tension experiments are given in Supplementary Tables 1 and 2, respectively. The dashed line ('Isotropic') is the predicted strength considering grain-size evolution but no fabric development, whereas the dotted line ('Torsion') is the predicted strength incorporating both grain-size evolution

and fabric development⁹. The solid black line ('Tension') is calculated using an evolution equation for the effect of crystallographic fabric on deformation (see Supplementary Information). Pole figures indicate grain orientations at the end of a particular torsion test. Pole figures (circles outside main panel) are equal-area, lower-hemisphere projections. The strength of the CPO is quantified with the M index²⁰. The number of grains plotted in pole figures are given ('no.'). Grey scale in pole figures indicates the density of data points as multiples of a uniform distribution. Arrows indicate the direction of shear.

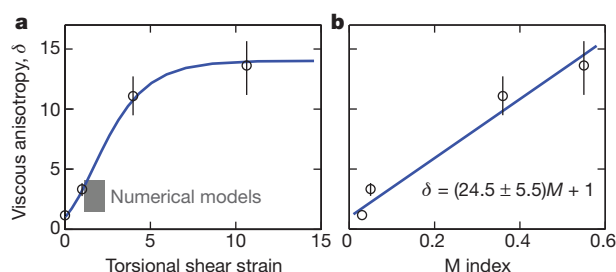


Figure 3 | Magnitude of viscous anisotropy, δ , as a function of strain and crystallographic-fabric strength. Error bars, 1 s.d. **a**, δ as a function of torsional shear strain. Circles denote values of δ calculated at a fixed stress from our torsion and tension data (see Supplementary Information). The blue line denotes δ versus shear strain calculated using evolution equations for the effect of crystallographic fabric on deformation (see Supplementary Information). The grey box denotes the range of δ predicted by numerical models^{7,8}. The range of torsional shear strains necessary to produce the starting fabrics used in the numerical models was estimated by comparing the reported fabric strengths to those observed in numerical¹² and laboratory¹⁶ experiments. **b**, δ as a function of the M index, a measure of the crystallographic fabric strength. This index ranges from 0 (for a random distribution of grain orientations) to 1 (for a perfect single crystal)²⁰. The blue line is a linear fit to the data (equation on graph) constrained to intercept the vertical axis at $\delta = 1$.

maximize the shear stress applied to the (010)[100] slip system relative to (2) the viscosity of olivine single crystals oriented such that negligible shear stress was applied to any slip system.

After reaching a peak at relatively low strain of ~ 0.25 , the stress measured during torsional deformation decreased significantly with increasing strain, approaching a steady-state value at a shear strain of ~ 5 (Fig. 2). Deformation was accompanied by both crystallographic fabric development and grain-size reduction. A previous study utilizing similar torsion experiments quantified the relative contributions of grain-size reduction and CPO development to weakening⁹, which allows us to compare the predicted stress–strain curve for olivine undergoing both grain-size reduction and CPO development (dashed line in Fig. 2) with that for olivine undergoing only grain-size reduction (dotted line in Fig. 2).

As previously observed for aggregates of San Carlos olivine^{15,16,18}, the degree of alignment of crystallographic axes increases with increasing strain (Fig. 2). By moderate strain (~ 4), [100] axes have aligned with the shear direction and [010] and [001] axes have gathered in a plane normal to the shear direction. Based on previous experiments (shear strain of ≤ 5), this crystallographic alignment was interpreted as the steady-state fabric characteristic of deformation at high stresses¹⁹. However, we demonstrate that this alignment is transient, as [010] and [001] axes gather into tight clusters by a shear strain of ~ 10 .

Stresses measured in torsion towards the end of a torsion test are significantly smaller than those measured in tension following the torsion test (Fig. 2). Samples are not only stronger in this tension test than predicted for an isotropic material—their strengths are greater than the peak stresses measured in the torsion test. Furthermore, stresses measured in tension increase with increasing torsional strain, reaching a maximum of ~ 215 MPa for torsional strains of ~ 10 . A tension test conducted before torsional deformation (red square in Fig. 2) confirms that an isotropic sample has the same strength during deformation in both torsion and tension.

The magnitude of δ after a particular amount of torsional strain can be obtained by comparing the stress measured in torsion to that subsequently measured in tension (for more detail, see Methods and Supplementary Information). Note that the tension and torsion data are compared after the same amount of torsional strain and, therefore, at the same grain size. Most of the increase in anisotropy occurs by a shear strain of ~ 10 (Fig. 3). The maximum value of $\delta \approx 14$ is about an order of magnitude larger than predicted by grain-scale numerical models^{7,8}, which is at least partly due to the numerical experiments

being limited to relatively weak fabrics. To better compare the results of numerical experiments to those presented here, the initial shear strain necessary to produce the simulated fabrics was estimated from the fabric strength. The comparison illustrates that the δ predicted by numerical models (grey box in Fig. 3) is in relatively good agreement with the evolution model presented here. This maximum value is still orders of magnitude smaller than the maximum VA of a single crystal⁵.

The value of δ at a particular shear strain correlates well with the strength of the crystallographic fabric measured with the M index after the same strain (Fig. 3b; the M index is described in Fig. 3 legend). Values of the M index can therefore be used as a proxy for δ . To our knowledge, the maximum M index in this study, $M = 0.55$, is the strongest measured for any experimentally or naturally deformed olivine aggregate. Thus, our measurements provide constraints on the maximum possible VA in the upper mantle. Additionally, because of the strong correlation between crystallographic fabric strength and seismic anisotropy^{16,20}, our results suggest that the magnitude of seismic anisotropy can also be used as a proxy for δ .

Previous studies identified dislocation-accommodated grain-boundary sliding as the dominant deformation mechanism in olivine-rich samples deformed at similar conditions to those used in this study. However, uncertainty still remains as to how, if at all, sliding on grain boundaries contributes to the macroscopic anisotropy. Although grain-shape preferred orientations²¹ have been shown to increase VA in metals^{22,23}, there is no clear correlation between grain-shape anisotropy measured in our samples and the magnitude of VA (see Supplementary Information). Thus, we suggest that the bulk of the VA results from strong crystallographic fabrics, which increase (or reduce) the shear stress in torsion (or tension) resolved onto the slip systems of intragranular dislocations accommodating grain-boundary sliding. This conclusion implies that, for a given CPO, VA similar to that measured here will also be present in materials deforming dominantly by other dislocation-related processes. For example, dislocation creep in which dislocation recovery is primarily mediated by dynamic recrystallization²⁴ should also result in a highly anisotropic viscosity.

Although many previous geodynamic studies did not include the VA of the upper mantle, our results demonstrate that significant anisotropy is developed by a relatively small (geologically speaking) amount of strain. Simulations of simple geodynamic systems, such as Couette flow away from a mid-ocean ridge, indicate that the upper mantle can be subject to shear strains of >60 (refs 25, 26). Using our results to calculate the magnitude of VA resulting from the shear strain in such a model, we suggest that most of the lithosphere and sublithospheric low-viscosity zone have values of δ near the predicted maximum. Although there is much complexity in extrapolating laboratory data to the deep upper mantle²⁷, a substantial portion of the upper mantle should be viscously anisotropic, such that processes involving both normal and shear deformation will be significantly affected.

The experiments described here demonstrate a marked difference between viscosities measured during normal and shear deformation and fully describe a two-dimensional, transversely isotropic viscosity. However, the crystallographic fabrics developed at the highest strain are no longer transversely isotropic (Fig. 2). Thus, geodynamic processes that involve three-dimensional flows or changing shear directions^{13,28} will need to consider other components of the viscosity tensor. Additionally, because the orientation and evolution of the viscosity tensor probably depends on the dominant slip system, future efforts to map a full viscosity tensor for textured olivine aggregates will also need to consider the effects of thermodynamic and environmental variables, such as pressure²⁹ and water content, on the VA of textured olivine aggregates.

METHODS SUMMARY

Samples of polycrystalline olivine with a composition (Fo_{50}) between that of forsterite (Fo) and fayalite (Fa) were fabricated by first reacting mixtures of Fe_2O_3 and SiO_2 powders in a gas-mixing furnace to form Fo_0 . The Fo_0 powders were then

mechanically mixed and reacted with San Carlos olivine (Fo₉₀) powders in the appropriate ratio to produce Fo₅₀. The resulting powders were cold pressed at a uniaxial pressure of 100 MPa into Ni cans with a central Ni post. These assemblies were then hot-pressed at a temperature of 1,200 °C and a confining pressure of 250 MPa for 8 h in an internally heated gas-medium apparatus^{17,30}.

Torsion experiments were performed by loading the sample into an Fe jacket along with alumina and zirconia pistons. Jacketed assemblies were loaded into a servo-controlled, internally heated, gas-medium apparatus fitted with a torsion actuator¹⁷. Torsion experiments were conducted at a constant equivalent strain rate of $2 \times 10^{-4} \text{ s}^{-1}$, a constant temperature of 1,200 °C, and a constant confining pressure of 250 MPa. Samples were removed from the apparatus, re-jacketed and reloaded into the apparatus for deformation in tension. Tension experiments were carried out under constant axial load at 1,200 °C and 250 MPa. Load stepping tests were performed by holding a single axial load until the strain rate reached an approximately constant value. After each load step, the load was returned to the value of the initial load step. To compare mechanical data from torsion and tension experiments, the values of the shear stress and shear strain rate measured in torsion were converted to equivalent (von Mises) values.

Sections were removed from the end of each sample for microstructural analysis after hot-pressing, torsional deformation and tensional deformation. Grain morphologies and CPOs were measured using electron-backscatter diffraction data collected in a JEOL 6500 scanning electron microscope.

Full Methods and any associated references are available in the online version of the paper.

Received 20 July; accepted 8 October 2012.

- Christensen, U. R. Some geodynamical effects of anisotropic viscosity. *Geophys. J. R. Astron. Soc.* **91**, 711–736 (1987).
- Han, D. & Wahr, J. An analysis of anisotropic mantle viscosity, and its possible effects on post-glacial rebound. *Phys. Earth Planet. Inter.* **102**, 33–50 (1997).
- Lev, E. & Hager, B. H. Rayleigh–Taylor instabilities with anisotropic lithospheric viscosity. *Geophys. J. Int.* **173**, 806–814 (2008).
- Lev, E. & Hager, B. H. Anisotropic viscosity changes subduction zone thermal structure. *Geophys. Geochem. Geosyst.* **12**, Q04009 (2011).
- Durham, W. B., & Goetze, C. Plastic flow of oriented single crystals of olivine 1. Mechanical data. *J. Geophys. Res.* **82**, 5737–5753 (1977).
- Castelnau, O., Blackman, D. K., Lebensohn, R. & Castañeda, P. P. Micromechanical modeling of the viscoplastic behavior of olivine. *J. Geophys. Res.* **113**, B09202, <http://dx.doi.org/10.1029/2007JB005444> (2008).
- Knoll, M., Tommasi, A., Logé, R. E. & Signorelli, J. W. A multiscale approach to model the anisotropic deformation of lithospheric plates. *Geochem. Geophys. Geosyst.* **10**, Q08009 (2009).
- Tommasi, A. *et al.* Structural reactivation in plate tectonics controlled by olivine crystal anisotropy. *Nature Geosci.* **2**, 423–427 (2009).
- Hansen, L. N., Zimmerman, M. E. & Kohlstedt, D. L. The influence of microstructure on deformation of olivine in the grain-boundary sliding regime. *J. Geophys. Res.* **117**, B09201, <http://dx.doi.org/10.1029/2012JB009305> (2012).
- Honda, S. Strong anisotropic flow in a finely layered asthenosphere. *Geophys. Res. Lett.* **13**, 1454–1457 (1986).
- Vauchez, A., Tommasi, A. & Barruol, G. Rheological heterogeneity, mechanical anisotropy and deformation of the continental lithosphere. *Tectonophysics* **296**, 61–86 (1998).
- Tommasi, A., Mainprice, D., Canova, G. & Chastel, Y. Viscoplastic self-consistent and equilibrium-based modeling of olivine lattice preferred orientations: implications for the upper mantle seismic anisotropy. *J. Geophys. Res.* **105**, 7893–7908 (2000).
- Castelnau, O., Blackman, D. K. & Becker, T. W. Numerical simulations of texture development and associated rheological anisotropy in regions of complex mantle flow. *Geophys. Res. Lett.* **36**, L12304, <http://dx.doi.org/10.1029/2009GL038027> (2009).
- Hansen, L. N., Zimmerman, M. E. & Kohlstedt, D. L. Grain-boundary sliding in San Carlos olivine: flow-law parameters and crystallographic-preferred orientation. *J. Geophys. Res.* **116**, B08201, <http://dx.doi.org/10.1029/2011JB008220> (2011).
- Zhang, S. & Karato, S. Lattice preferred orientation of olivine aggregates deformed in simple shear. *Nature* **375**, 774–777 (1995).
- Bystricky, M., Kunze, K., Burlini, L. & Burg, J. P. High shear strain of olivine aggregates; rheological and seismic consequences. *Science* **290**, 1564–1567 (2000).
- Paterson, M. S. & Olgaard, D. L. Rock deformation tests to large shear strains in torsion. *J. Struct. Geol.* **22**, 1341–1358 (2000).
- Zhang, S., Karato, S., Fitzgerald, J., Faul, U. H. & Zhou, Y. Simple shear deformation of olivine aggregates. *Tectonophysics* **316**, 133–152 (2000).
- Karato, S., Jung, H., Katayama, I. & Skemer, P. Geodynamic significance of seismic anisotropy of the upper mantle: new insights from laboratory studies. *Annu. Rev. Earth Planet. Sci.* **36**, 59–95 (2008).
- Skemer, P., Katayama, I., Jiang, Z. & Karato, S. The misorientation index: development of a new method for calculating the strength of lattice-preferred orientation. *Tectonophysics* **411**, 157–167 (2005).
- Launeau, P. & Cruden, A. R. Magmatic fabric acquisition mechanisms in a syenite; results of a combined anisotropy of magnetic susceptibility and image analysis study. *J. Geophys. Res.* **103**, 5067–5089 (1998).
- Bricknell, R. H. & Edington, J. W. Mechanical anisotropy and deformation mechanisms in an Al–Cu–Zr superplastic alloy. *Acta Metall.* **27**, 1313–1318 (1979).
- Barnett, M. R., Ghaderi, A., Sabirov, I. & Hutchinson, B. Role of grain boundary sliding in the anisotropy of magnesium alloys. *Scr. Mater.* **61**, 277–280 (2009).
- Platt, J. P. & Behr, W. M. Grain size evolution in ductile shear zones: implications for strain localization and the strength of the lithosphere. *J. Struct. Geol.* **33**, 537–550 (2011).
- Podolefsky, N. S., Zhong, S. & McNamara, A. K. The anisotropic and rheological structure of the oceanic upper mantle from a simple model of plate shear. *Geophys. J. Int.* **158**, 287–296 (2004).
- Behn, M. D., Hirth, G. & Elsenbeck, J. R. Implications of grain size evolution on the seismic structure of the oceanic upper mantle. *Earth Planet. Sci. Lett.* **282**, 178–189 (2009).
- Cordier, P., Amodeo, J. & Carrez, P. Modelling the rheology of MgO under Earth's mantle pressure, temperature and strain rates. *Nature* **481**, 177–180 (2012).
- Becker, T. W. & Kawakatsu, H. On the role of anisotropic viscosity for plate-scale flow. *Geophys. Res. Lett.* **38**, L17307, <http://dx.doi.org/10.1029/2011GL048584> (2011).
- Raterron, P., Girard, J. & Chen, J. Activities of olivine slip systems in the upper mantle. *Phys. Earth Planet. Inter.* **200–201**, 105–112 (2012).
- Paterson, M. S. in *The Brittle Ductile Transition in Rocks* (eds Duba, A. G., Durham, W. B., Handin, J. W. & Wang, H. F.) 187–194 (Geophys. Monogr. Ser. Vol. 56, American Geophysical Union, 1990).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by NSF grant EAR1214876. Parts of this work were carried out in the Institute of Technology Characterization Facility, University of Minnesota, which receives partial support from NSF through the NNIN programme. We thank T. Becker, B. Holtzman and A. Tommasi for discussions.

Author Contributions All authors discussed experiment design. L.N.H. conducted sample preparation, deformation experiments, and microstructural and mechanical analyses. M.E.Z. aided in sample preparation and deformation experiments. L.N.H. wrote the initial draft of the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.N.H. (lars.norman.hansen@gmail.com).

METHODS

Sample preparation. Samples of polycrystalline olivine with a composition intermediate to forsterite and fayalite (Fo₅₀) were used for this study. We chose this composition because an increase in the iron content of olivine increases the grain-growth kinetics producing a coarse-grained starting material and decreases the samples strength such that high strains could be reached in a reasonable amount of time. Powders were synthesized by first reacting mixtures of Fe₂O₃ and SiO₂ powders in a gas-mixing furnace to form Fo₀. The Fo₀ powders were then mechanically mixed and reacted with San Carlos olivine (Fo₉₀) powders in the appropriate ratio to produce Fo₅₀. To prepare thin-walled cylinders of polycrystalline olivine, the resulting powders were cold pressed at a uniaxial pressure of 100 MPa into Ni cans with a central Ni post. These assemblies were then hot-pressed at a temperature of 1,200 °C and a confining pressure of 250 MPa for 8 h in an internally heated gas-medium apparatus^{17,30}. The resulting thin-walled cylinders were approximately 11.5 mm in outer diameter, 7.5 mm in inner diameter, and 25 mm in length. To buffer the oxygen fugacity at Ni/NiO, samples remained jacketed in Ni with a central Ni post. Further details of the sample fabrication process can be found in ref. 9.

Deformation experiments. Torsion experiments were performed by inserting the sample into an Fe jacket along with alumina and zirconia pistons. Jacketed assemblies were loaded into a servo-controlled, internally heated, gas-medium apparatus fitted with a torsion actuator¹⁷. Torsion experiments were conducted at a constant equivalent strain rate of $2 \times 10^{-4} \text{ s}^{-1}$, a constant temperature of 1,200 °C, and a constant confining pressure of 250 MPa. Deformation was stopped once the desired shear strain was reached. Results from torsion experiments are given in Supplementary Table 1. Samples were removed from the apparatus, re-jacketed and reloaded into the apparatus for deformation in tension. Tension experiments were carried out under constant axial load at 1,200 °C and 250 MPa. Load stepping tests were performed by holding a single axial load until the strain rate reached an approximately constant value, which typically occurred by an axial strain of 2%. The load was then either increased or decreased and held constant in a similar manner. After each load step, the load was returned to the value of the initial load step to (1) ensure grain-size evolution was limited and (2) assess if any weakening was taking place due to necking of the sample. Results from tension experiments are given in Supplementary Table 2. To compare mechanical data from torsion and tension experiments, the values of the shear stress and shear strain rate measured in torsion were converted to equivalent (von Mises) values. Further details of these processes are given in Supplementary Information.

Sections were removed from the end of each sample for microstructural analysis subsequent to hot-pressing, torsional deformation and tensional deformation. Grain morphologies and crystallographic preferred orientations were measured using electron-backscatter diffraction (EBSD) data collected in a JEOL 6500 scanning electron microscope. EBSD maps were constructed by rastering the electron beam across the sample with a step size of 0.2–0.5 µm. Grain-boundary maps were calculated by defining grain boundaries to have misorientation angles $>10^\circ$.

Data normalization. The full data set is presented in Supplementary Tables 1 and 2 and plotted in Supplementary Fig. 1 before normalization of data from tension experiments to constant strain rate as presented in Fig. 2. The stress at the end of

torsion experiments decreases with increasing torsional strain and is always smaller than the value obtained for an isotropic sample (dashed line in Fig. 2). Strengths measured in tension are larger than the values obtained for isotropic samples. A fit to data from tension experiments (solid black line in Fig. 2) indicates a general increase in tensional strength with increasing torsional strain. To produce Fig. 2, all tension data were normalized to a strain rate of $\dot{\epsilon} = 2 \times 10^{-4} \text{ s}^{-1}$ using a value for the stress exponent of $n = 4.1$ and the steady-state relationship⁹ $\dot{\epsilon} \propto \sigma^n$.

Calculation of δ as a function of strain. Stress–strain data from torsion experiments on aggregates of olivine have been described by⁹

$$\dot{\epsilon}_s = \frac{A}{d^p} \left(\frac{\sigma_s}{F_s} \right)^n \quad (2)$$

where A is a material-dependent constant, d is the grain size, p is the grain-size exponent, F_s is a measure of the geometric softening due to fabric development, and n is the stress exponent. The rate of change of F_s is described by

$$\dot{F}_s = \frac{\dot{\epsilon}_s (F_s^{\text{ss}} - F_s)}{\epsilon_c} \quad (3)$$

where F_s^{ss} is the steady-state value of F_s and ϵ_c is a critical strain for microstructural evolution and is of the order of unity. A previous study has determined $F_s = 0.73$ for aggregates of Fo₅₀ olivine⁹. A similar approach is taken here to describe axial deformation normal to the torsional shear plane and yields

$$\dot{\epsilon}_n = \frac{A}{d^p} \left(\frac{\sigma_n}{F_n} \right)^n \quad (4)$$

where F_n is a measure of the geometric hardening due to fabric development. Thus, the rate of change of F_n is described by

$$\dot{F}_n = \frac{\dot{\epsilon}_n (F_n^{\text{ss}} - F_n)}{\epsilon_c} \quad (5)$$

Note that \dot{F}_n depends on $\dot{\epsilon}_s$ and not $\dot{\epsilon}_n$ because in this study we are only considering fabric developed in torsion and assume that strains during tensional deformation are small enough that any associated fabric development is negligible. The solid black line in Fig. 2 is a solution to equations (4) and (5). We find that a value of $F_n^{\text{ss}} = 1.39$ sufficiently fits the experimental data.

To calculate δ at a constant stress and as a function of strain, as presented in Fig. 3, equations (2) and (4) are used to calculate $\delta = \dot{\epsilon}_n / \dot{\epsilon}_s$, which yields

$$\delta = \left(\frac{F_n}{F_s} \right)^n \quad (6)$$

This formulation yields a maximum VA of $\delta = (F_n^{\text{ss}} / F_s^{\text{ss}})^n = (1.39 / 0.73)^{4.1} = 14$. To facilitate easy incorporation of δ into geodynamic simulations without needing to integrate a system of coupled differential equations, the evolution of δ with increasing shear strain, γ , can be reasonably described by

$$\delta = 13 \tanh(0.25\gamma) + 1 \quad (7)$$

Flickering gives early warning signals of a critical transition to a eutrophic lake state

Rong Wang^{1,2}, John A. Dearing¹, Peter G. Langdon¹, Enlou Zhang², Xiangdong Yang², Vasilis Dakos^{3,4} & Marten Scheffer³

There is a recognized need to anticipate tipping points, or critical transitions, in social–ecological systems^{1,2}. Studies of mathematical^{3–5} and experimental^{6–9} systems have shown that systems may ‘wobble’ before a critical transition. Such early warning signals¹⁰ may be due to the phenomenon of critical slowing down, which causes a system to recover slowly from small impacts, or to a flickering phenomenon, which causes a system to switch back and forth between alternative states in response to relatively large impacts. Such signals for transitions in social–ecological systems have rarely been observed¹¹, not the least because high-resolution time series are normally required. Here we combine empirical data from a lake-catchment system with a mathematical model and show that flickering can be detected from sparse data. We show how rising variance coupled to decreasing autocorrelation and skewness started 10–30 years before the transition to eutrophic lake conditions in both the empirical records and the model output, a finding that is consistent with flickering rather than critical slowing down^{4,12}. Our results suggest that if environmental regimes are sufficiently affected by large external impacts that flickering is induced, then early warning signals of transitions in modern social–ecological systems may be stronger, and hence easier to identify, than previously thought.

The concern that global social–ecological systems are adversely affected by the cumulative impact of multiple interacting drivers¹³ has spurred efforts to anticipate abrupt, nonlinear changes^{1,2,14–16}. As a result, there are increasing efforts to use system dynamical theory to identify early warning signals of critical transitions^{10,11}. For example, close to tipping points, the recovery rate from small perturbations becomes very slow¹⁰. In the natural fluctuations of a system, such critical slowing down may be signalled by rising levels of variance and autocorrelation^{3,4,10,17}. In addition, deformation of the basin of attraction before a critical transition may be detected from increasing skewness⁵. Such theoretical predictions are supported by studies of past climate change^{15,18} and manipulated lake ecosystems⁶ and by laboratory experiments with zooplankton⁷, phytoplankton⁸ and fungi⁹. But other studies of palaeoclimate records¹⁹ and ecological models²⁰ are less conclusive. One problem is that, typically, long time series of high-resolution data are needed. In addition, in stochastic systems with high levels of noise, shifts between alternative basins of attraction may occur far from the classical tipping points at which critical slowing down can be observed. Such situations can give rise to another phenomenon known as flickering, in which the system starts jumping back and forth between alternative basins of attraction^{10,11,21}. Here we use a multi-decadal time series from a lake together with model simulations and show that flickering is reflected in relatively low-resolution time series as bimodality and increasing variance, coupled to a decrease in autocorrelation and skewness.

We reconstruct and analyse historical changes in the Erhai Lake-catchment system in Yunnan, China (Supplementary Fig. 1). Monitoring data and official socio-economic statistics provide historical trends for lake water quality and lake water level, population density, land use

and climate from the 1950s to 2009 (Supplementary Information). Longer records of the lake ecosystem are based on laboratory analyses of three lake sediment cores, each of which contains the contemporary mud–water interface. Microscopic counts of fossil diatoms (siliceous algae) and chironomid (non-biting midge) head capsules give multi-decadal proxy records of the aquatic ecosystem to ~125 years ago, and in one core to ~750 years ago (Supplementary Information). Other sediment analyses give information about sediment provenance, organic matter and water chemistry. The timescales for the cores were obtained from a combination of ¹⁴C, ²¹⁰Pb and ¹³⁷Cs radionuclide determinations (Supplementary Information). We also use a simple model²² (Supplementary Information) to simulate phosphorus dynamics in a lake approaching eutrophication under a regime of strong external perturbations. The model is defined as follows:

$$dP = \left[\alpha - sP + r \frac{P^n}{P^n + 1^n} \right] dt + \sigma P dW \quad (1)$$

where P is phosphorus concentration, α is phosphorus input rate (the control parameter), r is the maximum recycling rate ($r = 1$), s is the phosphorus loss rate ($s = 1$), n is the strength of the recycling response to phosphorus concentrations ($n = 8$) and t is time. White noise is added through a Wiener process dW with scaling factor σ ($\sigma = 0.25$). We increased the phosphorus input rate, α , linearly in 2,000 time steps from 0.1 to 0.7, crossing the threshold at which the transition to eutrophication occurs at time step 1,848, when $\alpha = 0.6619$.

Microfossil and geochemical records (Fig. 1a–d) from dated lake sediment cores (Supplementary Figs 2 and 3) were used to reconstruct the trends in the state of lake diatom communities and water quality back to the 1880s, and these records seem to reproduce the abrupt change in algal states observed in recent monitored data, between 2001 and 2005 (Supplementary Fig. 4). From the combined monitored and lake sediment data, it seems that a profound transition in the algal community occurred around 2001. Historical records (Fig. 1e, f) of exogenous drivers (from 1950) strongly suggest that altered trends in algal community composition, starting in the 1960s, track with nutrient loading of the lake driven by agricultural intensification (Supplementary Information). Superimposed on nutrient loading, which is a ‘slow’ driving variable on a multi-decadal timescale, are the influences of ‘fast’ driving variables on annual and sub-annual timescales: these fast variables are short-lived changes in water volumes as a result of lake water-level regulation and low rainfall between 1980 and 2000, which together triggered temporary eutrophication events (Supplementary Information). We surmise that as aquatic productivity grew in response to increased nutrient concentrations (Fig. 1c, d), positive-feedback mechanisms (Supplementary Fig. 5) gradually strengthened the eutrophication process: oxygen depletion led to hypolimnetic anoxia (Fig. 1b) and to the recycling of biologically available phosphorus²² from the upper sediments (Fig. 1c). Despite a return to higher water levels in 2004–05 (Fig. 1e), the strengthened positive feedback (Supplementary Information) prevented recovery of the diatom communities as late as 2009.

¹Palaeoecological Laboratory, Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK. ²State Key Laboratory of Lake Science & Environment, Nanjing Institute of Geography & Limnology, Chinese Academy of Sciences, Nanjing 210008, China. ³Department of Aquatic Ecology and Water Quality Management, Wageningen University, PO Box 47, NL-6700 AA Wageningen, The Netherlands. ⁴Integrative Ecology Group, Estación Biológica de Doñana, CSIC, C/Américo Vespucio S/N, E-41092 Seville, Spain.

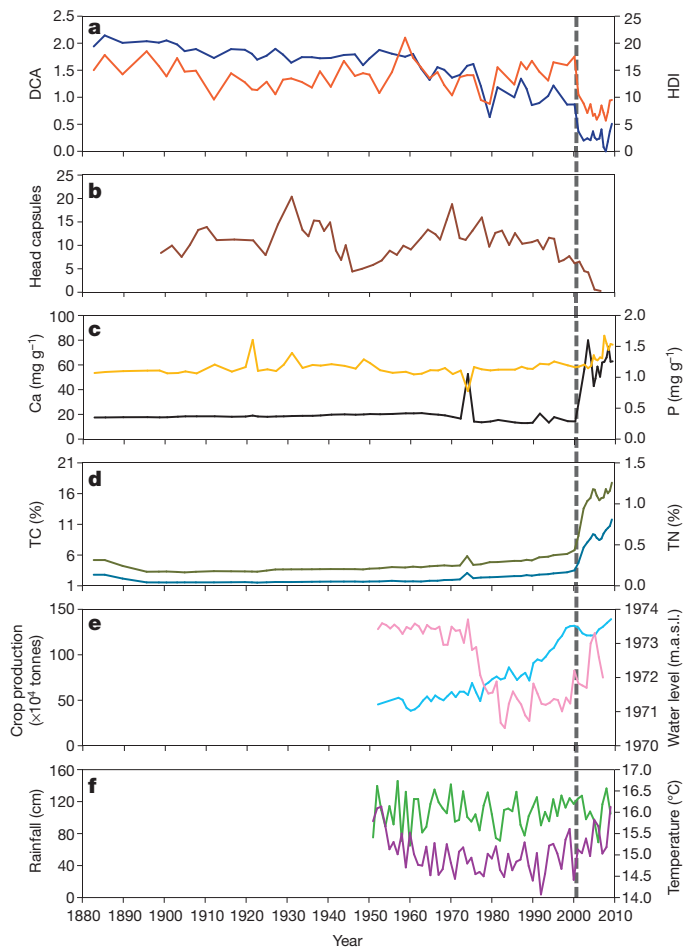


Figure 1 | Lake-sediment-based aquatic-system response variables and historical records of environmental drivers during the period 1883–2009. **a**, One-dimensional gradients of sediment diatom composition and diversity based on detrended correspondence analysis (DCA axis 1 scores) (blue) and calculated Hill's diversity index N2 (HDI) (orange) of non-interpolated sediment data. The DCA trend shows declining values from about 1965, with two sharp downward changes in about 1982 and about 2001 (dashed vertical line). The HDI data describe a more variable curve with a slight upward trend since the 1980s but an abrupt drop in about 2001. **b**, The abundance of chironomid (non-biting midge) larval head capsules in the sediment (number g^{-1} sediment) reflects the diverse lake conditions, and the very low abundance after 2001 indicates increasing anoxia in the hypolimnion²⁷. **c**, The abundance of calcium in the sediment (mg g^{-1} sediment) (black) rose in 1977 and after 2001, as a result of the biologically induced precipitation of CaCO_3 (calcite), which was linked to high rates of productivity by photosynthetic algae. Increased rates of photosynthesis decreased the level of dissolved CO_2 , raised the pH and produced local over-saturation of CaCO_3 . The total sediment phosphorus abundance (P mg g^{-1} sediment) (yellow) gradually rose from the 1960s; however, compared with calcium, there was only a small rise after 2003, suggesting partial depletion of sediment phosphorus as a result of anoxic recycling. **d**, Sediment total organic carbon (TC) (% of total sediment mass) (blue) and total nitrogen (TN) (% of total sediment mass) (green) gradually rose from the 1960s and abruptly rose after 2001, indicating rising levels of sedimenting organic matter as a result of increased aquatic productivity. **e**, Historical records of the annual crop yield in the region (Dali Bai, an autonomous prefecture of Yunnan) from a constant land area (blue) imply that agricultural intensification occurred from the 1960s and peaked initially in 2000. Historical records of the lowest annual lake water level (in metres above sea level (m.a.s.l.); pink), showing the combined effects of sluice building on outflow regulation for hydroelectric power and dry years after 1975. It should be noted that relatively high water levels were maintained in 2004 and 2005, after the trophic shift. **f**, Mean annual rainfall (green) and temperature (purple).

Several observations indicate¹¹ that the transition in Erhai Lake in ~2001 corresponds to the classic development of a bistable system. First, the shift in the state of the diatom communities (Fig. 1a) and the abrupt changes in other water quality indicators (Supplementary Fig. 4) are consistent with the behaviour of a lake that is shifting to an alternative, and stable, eutrophic state. Second, the observed multi-decadal increases in nutrient loading as a result of rising crop yields, together with the compounding effects of lake water level and climate variability (Fig. 1e, f), are consistent with the action of a 'slow', non-stationary driving variable that continues through the shift in diatom states. Third, the relatively low variability of diatom data over the past ~750 years suggests that system bistability is a recent phenomenon (Supplementary Fig. 6), and recent multi-decadal data show evidence for increasing bimodality (Fig. 2a). Fourth, statistical analyses and autoregression modelling allow rejection of null hypotheses for the absence of a critical transition and for the shift in diatom states caused by a shift in non-stationary driving variables (Fig. 2b, c and Supplementary Tables 1 and 2). Last, a phase-space plot for the diatom state response to varying concentrations of total dissolved phosphorus shows evidence of alternative states and hysteresis (Fig. 2d).

Detrended records of diatom composition and diversity indices (Fig. 3a, b) show a significant and continuously rising variance from ~1980 (Fig. 3c), whereas skewness (Fig. 3d) and autocorrelation (Fig. 3e) decline. Sensitivity tests (Supplementary Information and

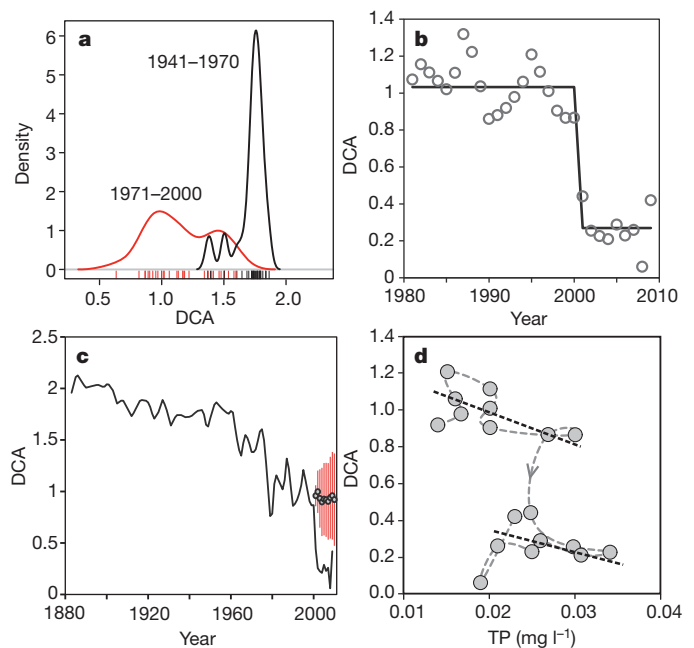


Figure 2 | Evidence for bistability, critical transition, alternative states and hysteresis in the DCA time series. **a**, Probability density functions (Gaussian kernel density estimation) for the periods 1941–70 (black) and 1971–2000 (red) show increasing bimodality, suggesting the development of a bistable system up to 2001. Short vertical bars show the density of individual points. **b**, A sequential analysis of DCA scores (circles) with mean values (horizontal line) for the period 1980–2010, using Student's *t*-test, shows a significant ($P \leq 0.01$) break point in about 2001. **c**, Predicted values (circles) in the period 2001–10 (vertical red bars) represent the probability levels ($P \leq 0.05$) using the optimum autoregressive integrated moving average (ARIMA) model (Supplementary Table 1) derived from the time series 1883–2000. The clear divergence between the observed values (solid black line) and the predicted values (circles) shows that the proposed transition at about 2001 cannot be explained in terms of the observed data. **d**, A phase-space plot of the dissolved nutrient driving variable (total phosphorus, TP) versus the diatom state response variable (DCA). The plot describes two linear clusters of points, 1992–2001 (upper black dashed line) and 2001–09 (lower black dashed line), suggesting two alternative diatom states for all TP values in the range 0.02–0.03 mg l^{-1} , which is equivalent to ~50% of the whole TP scale: this finding is strong evidence for alternative states and hysteresis.

Supplementary Fig. 7) indicate that these findings are robust. The declines in skewness and autocorrelation suggest that the increased variance cannot be explained by critical slowing down of a system close to equilibrium¹⁹. With strong evidence for exogenous drivers, we can reject an alternative explanation for increasing variance in terms of internal noise generated solely by endogenic changes¹⁹. Thus, the rising variance is most likely to represent the interaction of multiple exogenous drivers and the crossing of internal thresholds that magnify system responses and induce flickering between alternative attractors^{12,17}. In large data sets, flickering can be detected by distinct probability distributions of system states given by potential analysis^{23,24}. Our data are too sparse to carry out a full potential analysis, but evidence of flickering at Erhai Lake exists in the form of observed eutrophication events and algal blooms between 1980 and 2000 (Supplementary Information) and in the increased variance in the diatom indices (Fig. 3c). Flickering is also supported by the apparent bimodality¹⁷ in the frequency distribution of states (Fig. 2a). Overall, the system dynamics changed on a multi-decadal timescale, and the critical transition, in about 2001, was presaged by signals of rising variance caused by flickering that started 10–30 years previously (Fig. 3c and Supplementary Fig. 7i), a similar timescale to that found in mathematical models of lake regime shifts^{4,17}.

We checked whether the observed changes in autocorrelation, variance and skewness are consistent with flickering over a bistable region, by comparing the behaviour of these metrics to metrics estimated in

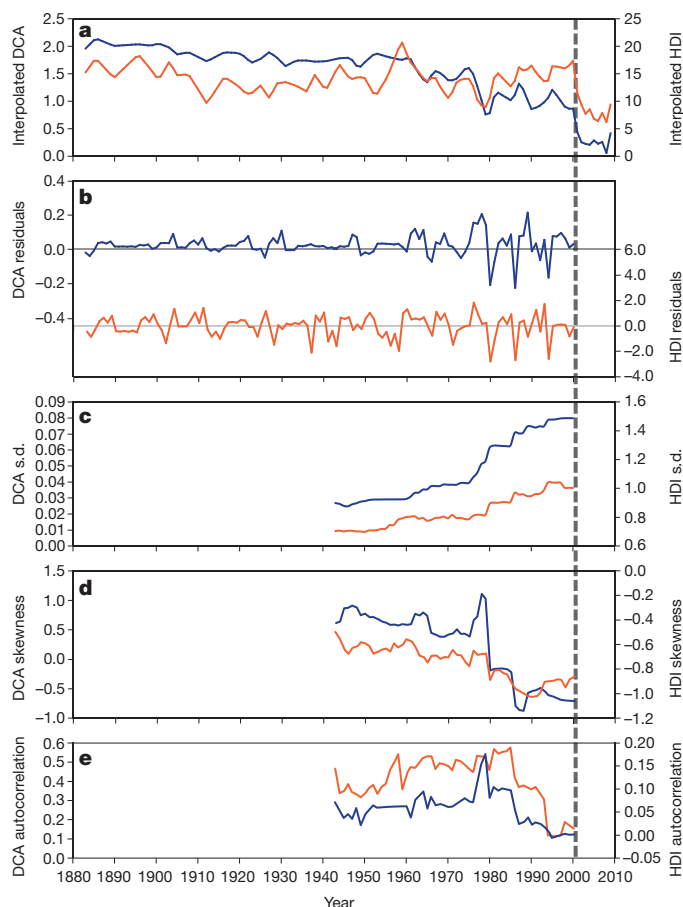


Figure 3 | Potential early warning signals of the regime shift in the lake trophic state for DCA (blue) and HDI (orange) time series. **a**, Interpolated one-dimensional gradients for sediment diatom composition (derived from Fig. 1a). **b**, Positive and negative residuals from **a**, with grey horizontal lines showing zero values. **c**, Variance of **b** using s.d. **d**, Skewness of **b**. **e**, Lag 1 autocorrelation of **b**. (**c**–**e**, Plots were calculated using a 59-year (half time series) sliding window through the period 1883–2001 and are plotted to the right of the window.) The dashed vertical line denotes 2001.

simulated time series from a simple model²² that describes the transition of a lake to eutrophic conditions (Supplementary Information). In this model, a positive feedback between phosphorus concentration and phosphorus recycling from the sediment causes alternative stable states (Fig. 4a). Under the strong noise regime that we impose, the

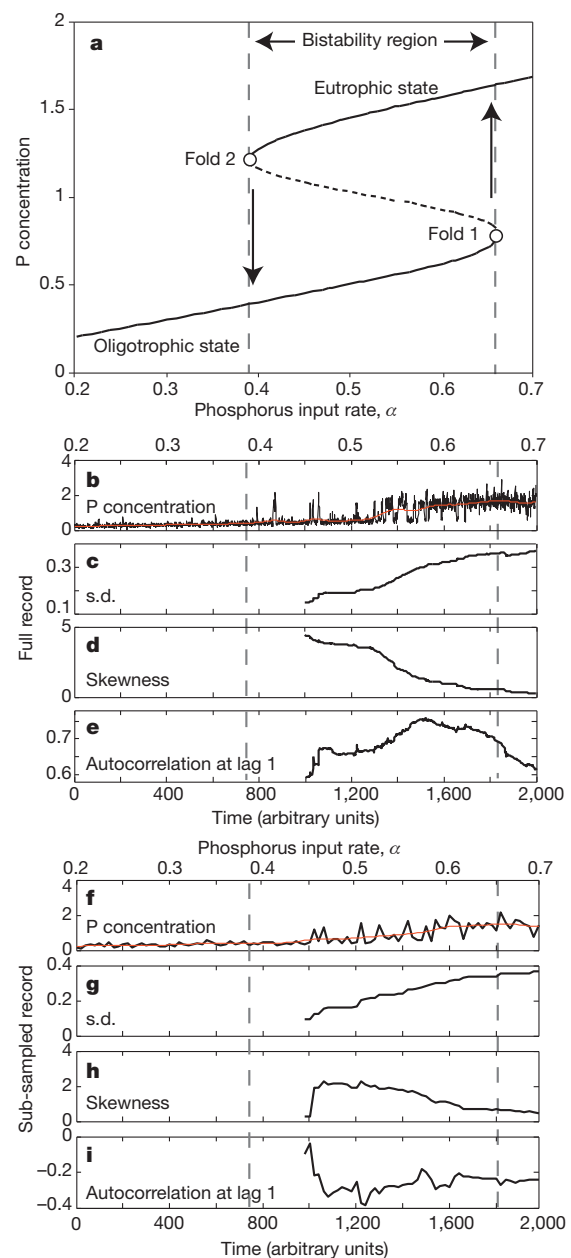


Figure 4 | Response of the lake model to increased phosphorus loading, and early warning signals. **a**, Equilibrium phosphorus concentration in the water as a function of the input rate. The middle section of the curve between the two bifurcation points, fold 1 and fold 2, consists of unstable saddle points that mark the border between the basins of attraction of the two stable branches of the hysteresis curve. **b**, Simulated response of the stochastically forced model to a gradual increase in phosphorus input rate over the same range covered by the axis in **a**. **c**–**e**, Early warning signals, s.d. (**c**), skewness (**d**) and lag 1 autocorrelation (**e**) for the full data, high-resolution, record (**b**). **f**, As for **b**, but sub-sampled (every 20 points) to give a coarse-resolution record. **g**–**i**, Early warning signals, s.d. (**g**), skewness (**h**) and lag 1 autocorrelation (**i**), for the coarse-resolution record (**f**). The vertical dashed lines mark the region of hysteresis in the non-stochastic model (**a**), corresponding to the parameter region with flickering in the stochastically forced model (**b**). Skewness, s.d. and autocorrelation were computed for a sliding window (a half time series) after removing a trend (red lines) using a Gaussian smoothing function and were plotted to the right of the window. All y-axis values are dimensionless.

dynamics over a range of parameter values are characterized by flickering between the alternative basins of attraction (Fig. 4b, f). This flickering results in bistability in the system before the system settles more permanently into an alternative state. Variance rises in both the original 'high-resolution' time series (Fig. 4c) and in a sub-sampled 'coarse-resolution' record (Fig. 4g). Skewness decreases to zero in both series (Fig. 4d, h). Autocorrelation rises in the high-resolution time series (Fig. 4e), until it drops in the last part of the record before the transition. But autocorrelation declines to zero or even negative values over the flickering region in the sub-sampled time series (Fig. 4i). Although published reports on systems with small fluctuations around equilibrium also show simultaneous increases in variance, autocorrelation and skewness¹¹, our results show that highly stochastic systems can produce contrasting patterns. The observed rise in variance, together with the simultaneous declines in autocorrelation and skewness, in the Lake Erhai records is thus in line with the findings expected for a flickering system sampled at relatively low time resolution.

Our findings have several implications for detecting transitions in real world systems. First, in systems in which exogenous drivers result in high levels of disturbance, flickering can be a more likely source of early warning signals than critical slowing down. Second, flickering will produce a rise in variance coupled to a decrease in skewness and autocorrelation at low time resolutions. Third, flickering can start before the stage at which the attractor of an alternative state becomes strong enough to capture the system from the noisy regime of forcing. Thus, a flickering signal in a modern system can be considered a direct warning that the system has left the 'safe operating space'¹⁶. Fourth, the potential advantages and ease of finding flickering signals suggests that every opportunity should be taken to study multi-decadal time series of modern social–ecological systems, even if the available data are of low resolution. Because monitored data do not always provide sufficiently long timescales of observations, especially for ecological variables, proxies such as the ones derived from sediment studies could be a useful resource in this respect^{25,26}.

Received 19 January; accepted 3 October 2012.

Published online 18 November 2012.

1. Carpenter, S. R. *et al.* Science for managing ecosystem services: beyond the Millennium Ecosystem Assessment. *Proc. Natl Acad. Sci. USA* **106**, 1305–1312 (2009).
2. Nicholson, E. *et al.* Priority research areas for ecosystem services in a changing world. *J. Appl. Ecol.* **46**, 1139–1144 (2009).
3. Brock, W. A. & Carpenter, S. R. Variance as a leading indicator of regime shift in ecosystem services. *Ecol. Soc.* **11**, 9 (2006).
4. Biggs, R., Carpenter, S. R. & Brock, W. A. Turning back from the brink: detecting an impending regime shift in time to avert it. *Proc. Natl Acad. Sci. USA* **106**, 826–831 (2009).
5. Guttal, V. & Jayaprakash, C. Changing skewness: an early warning signal of regime shifts in ecosystems. *Ecol. Lett.* **11**, 450–460 (2008).
6. Carpenter, S. R. *et al.* Early warnings of regime shifts: a whole-ecosystem experiment. *Science* **332**, 1079–1082 (2011).
7. Drake, J. M. & Griffen, B. D. Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459 (2010).
8. Veraart, A. J. *et al.* Recovery rates reflect distance to a tipping point in a living system. *Nature* **481**, 357–359 (2012).
9. Dai, L., Vorselen, D., Korolev, K. S. & Gore, J. Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* **336**, 1175–1177 (2012).
10. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
11. Scheffer, M. *Critical Transitions in Nature and Society* (Princeton Univ. Press, 2009).
12. Brock, W. A. & Carpenter, S. R. Interacting regime shifts in ecosystems: implication for early warnings. *Ecol. Monogr.* **80**, 353–367 (2010).
13. Beddington, J. Food security: contributions from science to a new and greener revolution. *Phil. Trans. R. Soc. B* **365**, 61–71 (2010).
14. International Council for Science. *ICSU Strategic Plan II 2012–2017* (International Council for Science, 2011).
15. Lenton, T. M. *et al.* Tipping elements in the Earth's climate system. *Proc. Natl Acad. Sci. USA* **105**, 1786–1793 (2008).
16. Rockström, J. *et al.* A safe operating space for humanity. *Nature* **461**, 472–475 (2009).
17. Carpenter, S. R. & Brock, W. A. Rising variance: a leading indicator of ecological transition. *Ecol. Lett.* **9**, 311–318 (2006).
18. Dakos, V. *et al.* Slowing down as an early warning signal for abrupt climate change. *Proc. Natl Acad. Sci. USA* **105**, 14308–14312 (2008).
19. Ditlevsen, P. D. & Johnsen, S. J. Tipping points: early warning and wishful thinking. *Geophys. Res. Lett.* **37**, L19703 (2010).
20. Hastings, A. & Wysham, D. B. Regime shifts in ecological systems can occur with no warning. *Ecol. Lett.* **13**, 464–472 (2010).
21. Taylor, K. C. *et al.* The 'flickering switch' of late Pleistocene climate change. *Nature* **361**, 432–436 (1993).
22. Carpenter, S. R., Ludwig, D. & Brock, W. A. Management of eutrophication for lakes subject to potentially irreversible change. *Ecol. Appl.* **9**, 751–771 (1999).
23. Livina, V. N., Kwasniok, F. & Lenton, T. M. Potential analysis reveals changing number of climate states during the last 60 kyr. *Clim. Past* **6**, 77–82 (2010).
24. Hirota, M., Holmgren, M., Van Nes, E. H. & Scheffer, S. Global resilience of tropical forest and savanna to critical transitions. *Science* **334**, 232–235 (2011).
25. Dearing, J. A., Braimoh, A. K., Reenberg, A., Turner, B. L. & van der Leeuw, S. Complex land systems: the need for long time perspectives to assess their future. *Ecol. Soc.* **15**, 21 (2010).
26. Dearing, J. A. *et al.* Extending the timescale and range of ecosystem services through paleoenvironmental analyses: the example of the lower Yangtze basin. *Proc. Natl Acad. Sci. USA* **109**, E1111–E1120 (2012).
27. Little, J. L., Hall, R. I., Quinlan, R. & Smol, J. P. Past trophic status and hypolimnetic anoxia during eutrophication and remediation of Gravenhurst Bay, Ontario: comparison of diatoms, chironomids, and historical records. *Can. J. Fish. Aquat. Sci.* **57**, 333–341 (2000).

Supplementary Information is available in the online version of the paper.

Acknowledgements A studentship to R.W. was funded through the UK Overseas Research Students Awards Scheme, the Great Britain–China Educational Trust and the University of Southampton. Laboratory analyses and fieldwork were supported by the National Basic Research Program of China (973 program, 2012CB956104). We are grateful to J. Shen and X. Chen for data collection, Y. L. Li for diatom analysis (core EH2), W. L. Xia for radionuclide measurements and Y. X. Zhu for geochemical analyses. V.D. is supported by an NWO-Rubicon and an EU Marie Curie fellowship. M.S. is funded by a Spinoza (NWO) Award and by a European Research Council Advanced Grant. We thank J. Dash and J. Noble for discussions about ARIMA models. This is a Sustainability Science at Southampton publication.

Author Contributions R.W. designed the research, undertook fieldwork, performed diatom and mathematical analyses, compiled historical data and wrote the paper. J.A.D. and P.G.L. designed and supervised the research and wrote the paper. E.Z. and X.Y. undertook fieldwork, performed and supervised other analyses, and commented on the manuscript. V.D. and M.S. carried out modelling experiments and wrote the paper. All authors discussed the results and commented on the final manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.A.D. (j.dearing@soton.ac.uk).

Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres

Andrew H. Paterson¹, Jonathan F. Wendel², Heidrun Gundlach³, Hui Guo¹, Jerry Jenkins^{4,5}, Dianchuan Jin⁶, Danny Llewellyn⁷, Kurtis C. Showmaker⁸, Shengqiang Shu⁴, Joshua Udall⁹, Mi-jeong Yoo², Robert Byers⁹, Wei Chen⁶, Adi Doron-Faigenboim¹⁰, Mary V. Duke¹¹, Lei Gong², Jane Grimwood^{4,5}, Corrinne Grover², Kara Grupp², Guanjin Hu², Tae-ho Lee¹, Jingping Li¹, Lifeng Lin¹, Tao Liu⁶, Barry S. Marler¹, Justin T. Page⁹, Alison W. Roberts¹², Elisson Romanel¹³, William S. Sanders⁸, Emmanuel Szadkowski², Xu Tan¹, Haibao Tang^{1,14}, Chunming Xu^{2,15}, Jinpeng Wang⁶, Zining Wang¹, Dong Zhang¹, Lan Zhang⁶, Hamid Ashrafi¹⁶, Frank Bedon⁷, John E. Bowers¹, Curt L. Brubaker^{7,17}, Peng W. Chee¹⁸, Sayan Das¹, Alan R. Gingle¹, Candace H. Haigler¹⁹, David Harker⁹, Lucia V. Hoffmann²⁰, Ran Hovav¹⁰, Donald C. Jones²¹, Cornelia Lemke¹, Shahid Mansoor^{1,22}, Mehboob ur Rahman²², Lisa N. Rainville¹, Aditi Rambani⁹, Umesh K. Reddy²³, Jun-kang Rong¹, Yehoshua Saranga²⁴, Brian E. Scheffler¹¹, Jodi A. Scheffler¹¹, David M. Stelly²⁵, Barbara A. Triplett²⁶, Allen Van Deynze¹⁶, Maite F. S. Vaslin²⁷, Vijay N. Waghmare²⁸, Sally A. Walford⁷, Robert J. Wright²⁹, Essam A. Zaki³⁰, Tianzhen Zhang³¹, Elizabeth S. Dennis⁷, Klaus F. X. Mayer³, Daniel G. Peterson⁸, Daniel S. Rokhsar⁴, Xiyin Wang^{1,6} & Jeremy Schmutz^{4,5}

Polyploidy often confers emergent properties, such as the higher fibre productivity and quality of tetraploid cottons than diploid cottons bred for the same environments¹. Here we show that an abrupt five- to sixfold ploidy increase approximately 60 million years (Myr) ago, and allopolyploidy reuniting divergent *Gossypium* genomes approximately 1–2 Myr ago², conferred about 30–36-fold duplication of ancestral angiosperm (flowering plant) genes in elite cottons (*Gossypium hirsutum* and *Gossypium barbadense*), genetic complexity equalled only by *Brassica*³ among sequenced angiosperms. Nascent fibre evolution, before allopolyploidy, is elucidated by comparison of spinnable-fibred *Gossypium herbaceum* A and non-spinnable *Gossypium longicalyx* F genomes to one another and the outgroup D genome of non-spinnable *Gossypium raimondii*. The sequence of a *G. hirsutum* A₁D₁ (in which ‘t’ indicates tetraploid) cultivar reveals many non-reciprocal DNA exchanges between subgenomes that may have contributed to phenotypic innovation and/or other emergent properties such as ecological adaptation by polyploids. Most DNA-level novelty in *G. hirsutum* recombines alleles from the D-genome progenitor native to its New World habitat and the Old World A-genome progenitor in which spinnable fibre evolved. Coordinated expression changes in proximal groups of functionally distinct genes, including a nuclear mitochondrial DNA block, may account for clusters of cotton-fibre quantitative trait loci affecting diverse traits. Opportunities abound for dissecting emergent properties of other polyploids, particularly angiosperms, by comparison to diploid progenitors and outgroups.

The *Gossypium* genus is ideal for investigating emergent consequences of polyploidy. A-genome diploids native to Africa and Mexican D-genome diploids diverged ~5–10 Myr ago⁴. They were

reunited ~1–2 Myr ago by trans-oceanic dispersal of a maternal A-genome propagule resembling *G. herbaceum* to the New World², hybridization with a native D-genome species resembling *G. raimondii*, and chromosome doubling (Fig. 1). The nascent A₁D₁ allopolyploid spread throughout the American tropics and subtropics, diverging into at least five species; two of these species (*G. hirsutum* and *G. barbadense*) were independently domesticated to spawn one of the world’s largest industries (textiles) and become a major oilseed.

New insight into *Gossypium* biology is offered by a genome sequence of *G. raimondii* Ulbr. (chromosome number, 13) with ~8× longer scaffold N50 (18.8 versus 2.3 megabases (Mb)) compared with a draft⁵, and oriented to 98.3% (versus 52.4%)⁵ of the genome (Supplementary Table 1.3a). Across 13 pseudomolecules totalling 737.8 Mb, ~350 Mb (47%) of euchromatin span a gene-rich 2,059 centimorgan (cM), and ~390 Mb (53%) of heterochromatin span a repeat-rich 186 cM (Supplementary Discussion, sections 1.5 and 2.1). Despite having the least-repetitive DNA of the eight *Gossypium* genome types, *G. raimondii* is 61% transposable-element-derived (Supplementary Table 2.1). Long-terminal-repeat retrotransposons (LTRs) account for 53% of *G. raimondii*, but only 3% of LTR base pairs derive from 2,345 full-length elements. The 37,505 genes and 77,267 protein-coding transcripts annotated (Supplementary Table 2.3 and <http://www.phytozome.com>) comprise 44.9 Mb (6%) of the genome, largely in distal chromosomal regions (Supplementary Discussion, section 2.1).

Shortly after its divergence from an ancestor shared with *Theobroma cacao* at least 60 Myr ago⁶, the cotton lineage experienced an abrupt five- to sixfold ploidy increase. Individual grape chromosome segments resembling ancestral eudicot genome structure, or corresponding cacao chromosome segments, generally have five

¹Plant Genome Mapping Laboratory, University of Georgia, Athens, Georgia 30602, USA. ²Department of Ecology, Evolution and Organismal Biology, Iowa State University, Ames, Iowa 50011, USA. ³MIPS/IBIS Institute for Bioinformatics and System Biology, German Research Center for Environmental Health (GmbH), 85764 Neuherberg, Germany. ⁴Department of Energy Joint Genome Institute, Walnut Creek, California 94595, USA. ⁵HudsonAlpha Institute of Biotechnology, Huntsville, Alabama 35806, USA. ⁶Center for Genomics and Computational Biology, School of Life Sciences, and School of Sciences, Hebei United University, Tangshan, Hebei 063000, China. ⁷CSIRO Plant Industry, Canberra, ACT 2601, Australia. ⁸Institute for Genomics, Biocomputing & Biotechnology, Mississippi State University, Mississippi State, Mississippi 39762, USA. ⁹Plant and Wildlife Science Department, Brigham Young University, Provo, Utah 84602, USA. ¹⁰Department of Field Crops, Plant Sciences Institute, ARO, Bet-Dagan 50250, Israel. ¹¹Jamie Whitten Delta States Research Center, USDA-ARS, Stoneville, Mississippi 38776, USA. ¹²Department of Biological Sciences, University of Rhode Island, Kingston, Rhode Island 02881, USA. ¹³Departamento de Genética, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 21941-901, Brazil. ¹⁴J. Craig Venter Institute, Rockville, Maryland 20850, USA. ¹⁵Key Laboratory of Molecular Epigenetics of MOE, Unit of Plant Epigenetics, Institute of Genetics & Cytology, Northeast Normal University, Renmin Street, 5268 Changchun, China. ¹⁶Plant Reproductive Biology Extension Center, University of California, Davis, California 95616, USA. ¹⁷Bayer CropScience, Technologiepark 38, 9052 Gent, Belgium. ¹⁸Coastal Plain Experiment Station, University of Georgia, Tifton, Georgia 31793, USA. ¹⁹Departments of Crop Science and Plant Biology, North Carolina State University, Raleigh, North Carolina 27695, USA. ²⁰Centro Nacional de Pesquisa em Algodão, EMBRAPA, Santo Antônio de Goiás, GO 75375-000, Brazil. ²¹Cotton Incorporated, Cary, North Carolina 27513, USA. ²²National Institute for Biotechnology & Genetic Engineering, Faisalabad 38000, Pakistan. ²³Department of Biology, West Virginia State University, Institute, West Virginia 25112, USA. ²⁴Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, The Hebrew University of Jerusalem, Rehovot 76100, Israel. ²⁵Department of Soil and Crop Science, Texas A&M University, College Station, Texas 77843, USA. ²⁶Cotton Fiber Bioscience Research, USDA-ARS, New Orleans, Louisiana 70124, USA. ²⁷Departamento de Microbiologia, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro 21941-971, Brazil. ²⁸Central Institute for Cotton Research, Nagpur, 440010 Maharashtra, India. ²⁹Department of Plant Sciences, Texas Tech University, Lubbock, Texas 79415, USA. ³⁰Nucleic Acids Department, Genetic Engineering & Biotechnology Research Institute, 21934 Alexandria, Egypt. ³¹Key Laboratory of Crop Genetics and Germplasm Enhancement, Nanjing Agricultural University, Nanjing, 210095 Jiangsu, China.

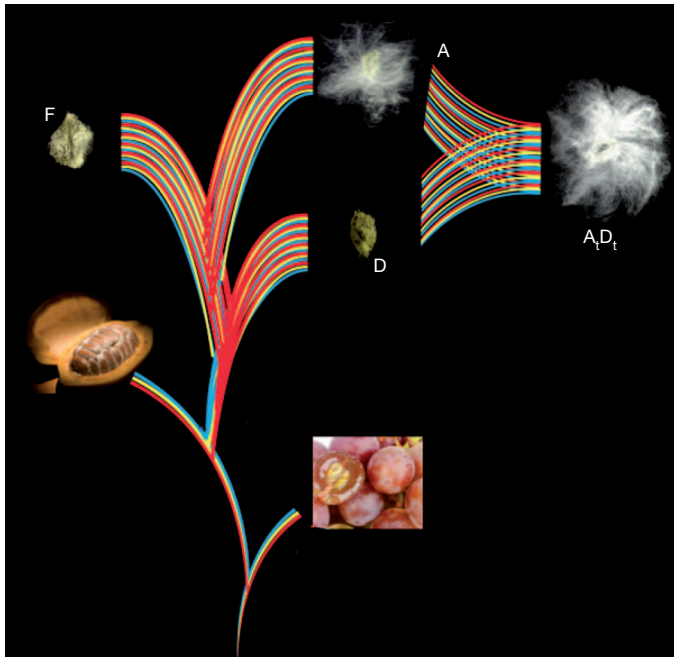


Figure 1 | Evolution of spinnable cotton fibres. Paleohexaploidy in a eudicot ancestor (red, yellow and blue lines) formed a genome resembling that of grape (bottom right). Shortly after divergence from cacao (bottom left), the *Gossypium* lineage experienced a five- to sixfold ploidly increase. Spinnable fibre evolved in the A genome after its divergence from the F genome, and was further elaborated after the merger of A and D genomes ~1–2 Myr ago, forming the common ancestor of *G. hirsutum* (Upland) and *G. barbadense* (Egyptian, Sea Island and Pima) cottons.

(infrequently six) best-matching *G. raimondii* regions and secondary matches resulting from pan-eudicot hexaploidy^{7,8} (Fig. 2 and Supplementary Table 3.1). Paralogous genes tracing to this five- to sixfold

ploidy increase show a single peak of synonymous nucleotide-substitution (K_s) values, suggesting either one, or multiple closely spaced, event(s) (Supplementary Fig. 3.5). Pairwise cytological similarity among A-genome chromosomes⁹ suggests the most recent event was a duplication.

Paleopolyploidy may have accelerated cotton mutation rates: for 7,021 co-linearity-supported gene triplets, K_s rates and non-synonymous nucleotide-substitution (K_a) rates were, respectively, 19% and 15% larger for cotton–grape than cacao–grape comparisons (Supplementary Table 3.2). Adjusted for this acceleration (Supplementary Fig. 3.5), the cotton ploidy increase occurred about halfway between the pan-eudicot hexaploidy (<125 Myr ago¹⁰) and the present, near the low end of an estimated range of 57–70 Myr ago¹¹.

Paleopolyploidy increased the complexity of a Malvaceae-specific clade of Myb family transcription factors, perhaps contributing to the differentiation of epidermal cells into fibres rather than the mucilages of other Malvaceae. Among 204 R2R3, 8 R1R2R3 and 194 heterogeneous Myb transcription factors in *G. raimondii* (Supplementary Table 3.5), subgroup 9 has six members known only in Malvaceae (Fig. 3a), comprising a possible ‘fibre clade’ distinct from the *Arabidopsis thaliana* GL1-like subgroup 15 involved in trichome and root hair initiation and development¹². Expressed predominantly in early fibre development, elite cultivated tetraploid cottons have higher expression of five (50%) of ten subgroup 9 genes compared with wild (undomesticated) tetraploids (Fig. 3a and Supplementary Table 5.3). Some subgroup 9 genes are also active in leaves, hypocotyls and cotyledons (Supplementary Fig. 3.8), consistent with specialization for different types of epidermal cell differentiation such as production of a ‘pulpy layer’ secreted from the teguments surrounding cacao seeds, and mucilages in other Malvaceae fruit (*Abelmoschus* (okra), *Cola* (kola)) and roots (*Althaea* (marshmallow)).

Cotton growers were early adopters of integrated pest management¹³ strategies to deploy intrinsic defences conferred by pest- and disease-resistance genes that evolved largely after the 5–6-fold ploidy increase.

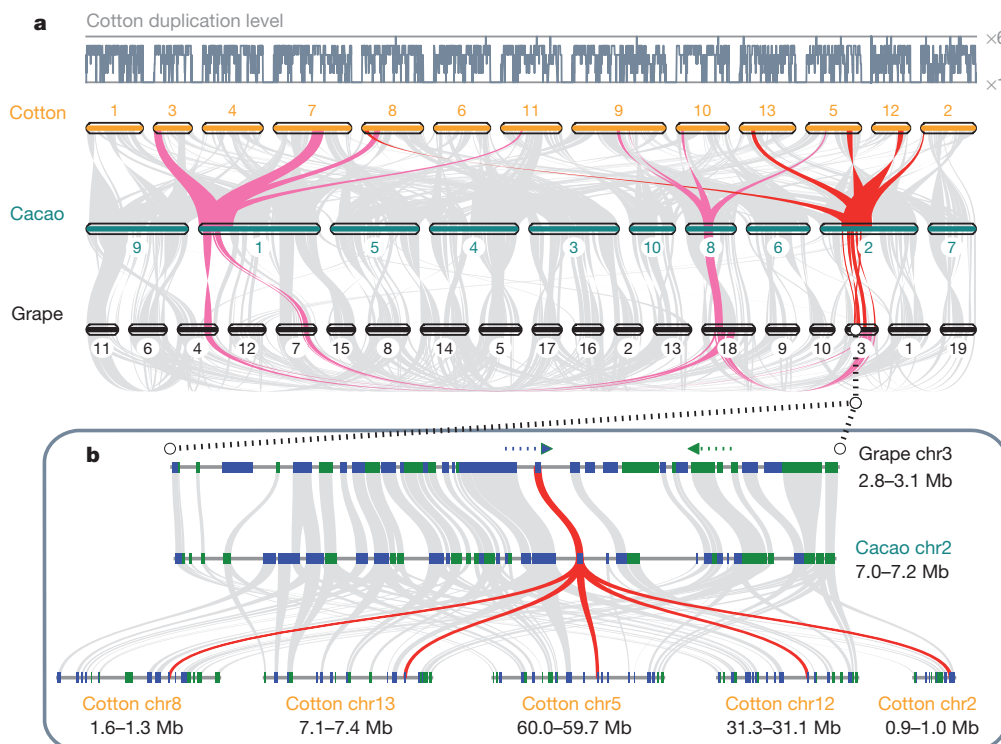


Figure 2 | Syntenic relationships among grape, cacao and cotton. **a**, Macro-syntenic connecting blocks of >30 genes (grey lines). Highlighted regions (pink and red) trace to a common ancestor before the pan-eudicot hexaploidy⁷, with the *Gossypium* lineage five- to sixfold ploidy increase forming multiple derived

regions. Inferred duplication depth in cotton varies (top). **b**, Micro-syntenic of grape chromosome (Chr) 3, cacao chromosome 2 and five cotton chromosomes. Rectangles represent predicted genes, with connecting grey lines showing co-linear relationships. An example (1 grape, 1 cacao, 5 cotton) is highlighted in red.

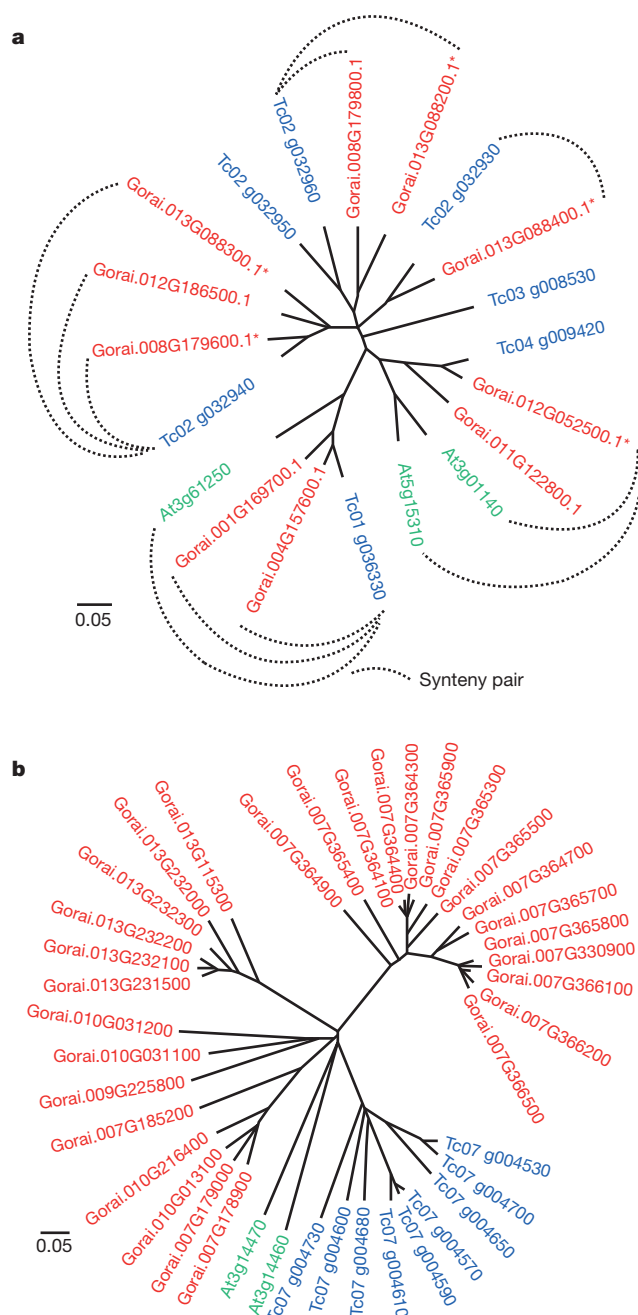


Figure 3 | Paleoevolution of cotton gene families. **a**, Myb subgroup 9 (ref. 12) originated from a gene on the progenitor of cacao chromosome 2 that formed two adjacent copies after Malvales–Brassicales divergence and then triplicated in cotton, with subsequent loss of one chromosome 8 and two chromosome 12 paralogues. One extant paralogue traces to pan-eudicot hexaploidy, Tc04 g009420, and reduplicated in cotton (Gorai.012G052500.1 and Gorai.011G122800.1) and *Arabidopsis*⁸ (At3g01140 and At5g15310). The other, Tc01 g036330, has reduplicated in cotton (Gorai.004G157600.1 and Gorai.001G169700.1). Asterisk indicates increased gene expression in elite versus wild tetraploids (Supplementary Table 5.3). **b**, The most NBS-rich region of *T. cacao*, on chromosome 7, corresponds to regions of *G. raimondii* chromosome triplets 2/10/13 and 7/9/4. Cacao chromosome 7 NBSs form a single branch, indicating lineage-specific expansion. *G. raimondii* chromosome 7 and 13 NBSs form distinct branches, indicating cluster/tandem duplication (gene numbers also reflect physical proximity of genes to one another).

A total of 300 (0.8%) *G. raimondii* genes encode nucleotide-binding site (NBS) domains (Supplementary Table 3.6), largely of coiled-coil (CC)-NBS and CC-NBS-leucine rich repeat subgroups (165, 55%). Like cereals¹⁴, after paleopolyploidy *G. raimondii* evolved clusters of new

NBS-encoding genes. The most NBS-rich (21%) region of *T. cacao*, on chromosome 7, corresponds to parts of *G. raimondii* chromosome triplets 2/10/13 and 7/9/4. In total, 27% and 25% of 294 mapped *G. raimondii* NBS genes are on these parts of chromosomes 7 and 9, often clustered in otherwise gene-poor surroundings (Supplementary Fig. 2.2). Most NBS clusters are species and chromosome specific (Fig. 3b and Supplementary Table 3.7), indicating rapid turnover and/or concerted evolution after cotton paleopolyploidy. In total, 230 (76.7%) NBS-encoding genes have experienced striking mutations (as detailed below) in the A genome since A–F divergence, reflecting an ongoing plant–pathogen ‘arms race’ (Supplementary Table 3.8).

Changes in gene expression during domestication have contributed to the deposition of >90% cellulose in cotton fibres, single-celled models for studying cell wall and cellulose biogenesis¹⁵. *G. raimondii* has at least 15 cellulose synthase (*CESA*) sequences required for cellulose synthesis¹⁶ (Supplementary Table 3.3), with four single-gene *Arabidopsis* clades having three (*CESA3*, required in expanding primary walls) or two (*CESA4*, *CESA7* and *CESA8*, each required in the thickening of secondary walls) clade members in *G. raimondii*¹⁶. *G. raimondii* has at least 35 cellulose-synthase-like (*CSL*) genes required for synthesis of cell wall matrix polysaccharides that surround cellulose microfibrils¹⁶ (Supplementary Table 3.4), including one family (*CSLJ*) absent in *Arabidopsis*¹⁶. Elite tetraploids have higher expression than wild cottons in 6 (40%) of 15 *CESA* genes and 12 (34%) of 35 *CSL* genes (Supplementary Table 5.3).

A total of 364 *G. raimondii* microRNA precursors from 28 conserved and 181 novel families (Supplementary Table 3.12), are predicted regulators of 859 genes enriched for molecule binding factors, catalytic enzymes, transporters and transcription factors (Supplementary Fig. 3.11, 12). Four conserved and 35 novel mRNAs were specifically expressed in *G. hirsutum* fibres, respectively targeting 53 and 318 genes, most with homology to proteins involved in fibre development (Supplementary Table 3.14, 15). Among 183,690 short interfering RNAs (siRNAs) found, 33,348 (18.15%) were on chromosome 13 (Supplementary Fig. 3.12), a vast enrichment. Small RNA^{17–19} biogenesis proteins include 13 argonaute, 6 dicer-like (*DCL*) and 5 RNA-dependent RNA polymerase orthologues (Supplementary Table 3.16). *G. raimondii* seems to be the first eudicot with two *DCL3* genes and two genes encoding RNA polymerase IVa (Supplementary Table 3.16), perhaps relating to control of its abundant retrotransposons.

From unremarkable hairs found on all *Gossypium* seeds, ‘spinnable’ fibres (fibres with a ribbon-like structure that allows for spinning into yarn) evolved in the A genome after divergence from the B, E and F genomes ~5–10 Myr ago⁴ (Fig. 1). To clarify the evolution of spinnable fibres, we sequenced the *G. herbaceum* A and *G. longicalyx* F genomes, which respectively differ from *G. raimondii* by 2,145,177 single-nucleotide variations (SNVs) and 477,309 indels, and 3,732,370 SNVs and 630,292 indels.

Specific genes are implicated in initial fibre evolution by both whole-gene and individual-nucleotide analyses. Across entire genes, 36 *G. herbaceum*–*G. raimondii* and 11 *G. herbaceum*–*G. longicalyx* orthologue pairs show evidence of diversifying selection ($\omega > 1$, $P < 0.05$) (Supplementary Table 4.1). A notable example, with *G. herbaceum*–*G. raimondii* $\omega > 9$, is Gorai.009G035800, a germin-like protein that is differentially expressed between normal and naked-seed cotton mutants during fibre expansion²⁰ and between wild and elite *G. barbadense* at 10 days post-anthesis (DPA; Supplementary Table 5.3).

Among 114,202 SNVs in 29,015 *G. herbaceum* genes after *G. herbaceum*–*G. longicalyx* A–F divergence (using D as outgroup, so F is the same as D, and A differs from both), we identified striking mutations including 1,090 non-synonymous mutations in 959 genes comprising the most severe 1% of functional impacts inferred using a modified entropy function²¹; 3,525 frameshift mutations (3,021 genes), 1,077 (987) premature stops, 527 (513) splice-site mutations, 102 (102) initiation alterations and 95 (94) extended reading frames (Supplementary Table 4.2, 3). These striking mutations have an average genomic

distribution (Supplementary Fig. 2.2) but are over-represented in genes coding for cell-wall-associated, kinase or nucleotide-binding proteins (Supplementary Table 4.5).

Striking mutations in the A-genome lineage are enriched ($P = 2.6 \times 10^{-18}$; Supplementary Discussion, section 4.4) within fibre-related quantitative trait locus (QTL) hotspots in A_tD_t tetraploid cottons²², suggesting that post-allopolyploidy elaboration of fibre development¹ involved recursive changes in A_t and new changes in D_t genes. Striking A-genome mutations have orthologues in 1,051 D_t and 951 A_t fibre QTL hotspots. Likewise, sequencing of *G. hirsutum* cultivar Acala Maxxa revealed 495 striking mutations in 391 genes, with 83 (21.2%) in D_t fibre QTL hotspots and 73 (18.7%) in A_t hotspots (Supplementary Table 4.6).

QTL hotspots affecting multiple fibre traits²² may reflect coordinated changes in expression of functionally diverse cotton genes. A total of 671 (1.79%) genes with >100 reads per million reads were differentially expressed in fibres from wild versus domesticated *G. hirsutum* (mostly at 10 DPA) and/or *G. barbadense* (mostly at 20 DPA) (Supplementary Table 5.3). Among 48 genes upregulated in domesticated *G. hirsutum* at 10 DPA, 20 (42%) are among 1,582 (4.2%) genes within QTL hotspot $D_t09.2$ (ref. 22) affecting length, uniformity, and short-fibre content, with 13 (27%) out of 677 (1.8%) genes in homeologous hotspot A_t09 affecting fibre elongation and fineness. Out of 45 genes downregulated in domesticated *G. barbadense* at 20 DPA, 16 (35.6%) map to $D_t09.2$, and 8 (17.7%) to A_t09 . In 79% of cultivated *G. barbadense*, this A_t region (which was then thought to be on chromosome 5, and is now known to be on chromosome 9) has been unconsciously introgressed by plant breeders with *G. hirsutum* DNA, suggesting an important contribution to productivity of *G. barbadense* cultivars²³.

A putative nuclear mitochondrial DNA (NUMT) sequence block²⁴ has an intriguing relationship with fibre improvement. A *G. raimondii* chromosome 1 region includes many genes closely resembling mitochondrial homologues ($K_s \sim 0.22$; Supplementary Table 4.7a). NUMT genes experienced a coordinated change in expression associated with *G. barbadense* domestication. The 105 (0.2%) genes upregulated in 10 DPA fibre of wild (versus elite) tetraploid *G. barbadense* (Supplementary Table 5.3) include 30 (37%; $P < 0.001$) of the 81 NUMT genes, including 8 NADH dehydrogenase and 4 cytochrome-c-related genes. All are within the QTL hotspot D_t01 that affects fibre fineness, length, and uniformity²², suggesting a fibre-specific change in electron transfer in *G. barbadense* domestication.

Emergent features of polyploids may be related to processes that render them no longer the sum of their progenitors and permit them to explore transgressive phenotypic innovations. Despite the A-genome origin of spinnable fibres, after 1–2 Myr of co-habitation in tetraploid nuclei most A_t and D_t homeologues are now expressed in fibres at similar levels (Supplementary Table 5.4). Such convergence is not ubiquitous: gene families involved in the synthesis of seed oil show strong A bias in wild *G. hirsutum* and its sister *G. tomentosum*, but strong D bias in an improved *G. hirsutum* (Supplementary Table 5.6).

Recruitment of D_t -genome genes into tetraploid fibre development¹ may have involved non-reciprocal DNA exchanges from A_t genes. In the ~40% of Acala Maxxa A_t and D_t genes that differ in sequence from their diploid progenitors (Fig. 4), most mutations are convergent, with A_t genes converted to the D_t state at more than twice the rate (25%) as the reciprocal (10.6%). Known to occur between cereal paralogues diverged by 70 Myr¹⁴, non-reciprocal DNA exchanges are more abundant between cotton A_t and D_t genes separated by only ~5–10 Myr⁴. Such non-reciprocal exchanges explain prior observations including incongruent gene tree topology for 10% (3 pairs) of *G. hirsutum* A_t and D_t homeologues in sequenced bacterial artificial chromosomes (BACs) (Supplementary Discussion, section 5.3); 13.2% of tetraploid DNA markers that showed different subgenomic affinities compared with the chromosomes to which they mapped, 9 of 13 being D_t biased (A_t to D_t)²⁵; and expressed-sequence-tag-based

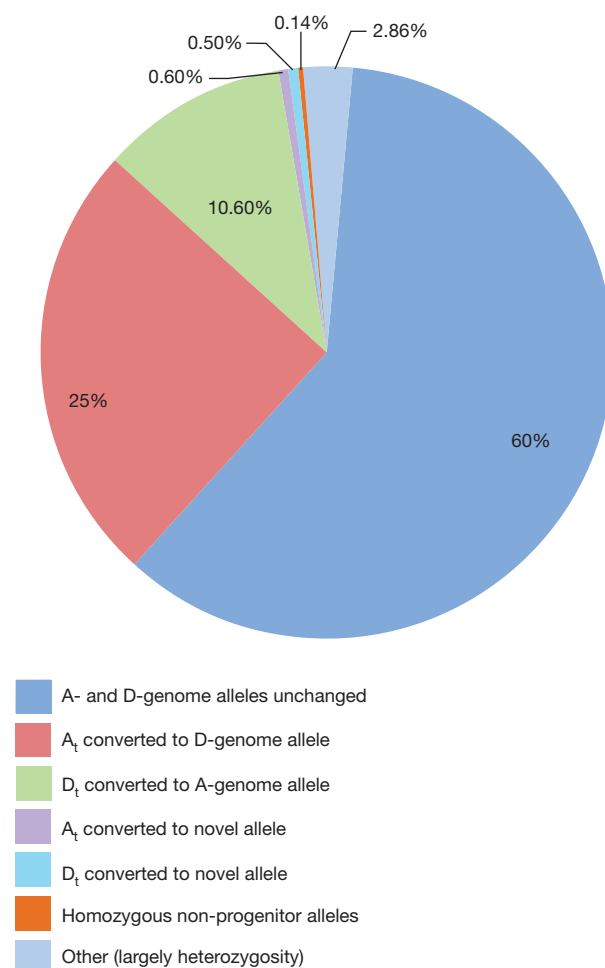


Figure 4 | Allelic changes between A- and D-genome diploid progenitors and the A_t and D_t subgenomes of *G. hirsutum* cultivar Acala Maxxa.

evidence of phylogenetic incongruity for as many as 7% of homeologous genes²⁶.

Several factors may have favoured D_t -biased allele conversion in tetraploid cotton. The nascent polyploid may have gained fitness from D-genome alleles native to its New World habitat. Before fortifying its reproductive barriers, the nascent polyploid may have occasionally outcrossed to nearby D-genome diploids, increasing the likelihood of illegitimate recombination. Outcrossing may also have contributed to the origin of *Gossypium gossypoides*, sister to *G. raimondii* and the only D-genome cotton containing many otherwise A-genome-specific repetitive DNAs^{27–29}. D_t -biased allele conversion may have contributed to slightly greater protein-coding nucleotide diversity in the A_t compared with the D_t -genome (Supplementary Table 5.7).

Whereas the *G. raimondii* reference sequence and *G. hirsutum* short-read sequences reveal much about tetraploid cotton genome structure and polyploid evolution, high-contiguity sequencing of polyploids may elucidate still-cryptic features. Tetraploid cotton sequencing appears feasible: among six pairs of A_t and D_t BAC clones, the most similar pair shows 99.1% shared D_t -D and 97.6% A_t -D content (Supplementary Table 6.1), sufficient divergence to de-convolute shotgun sequence to the correct subgenome. Increased knowledge of molecular diversity is a foundation for integrating genomics with ecological and field-level knowledge of *Gossypium* species and their diverse adaptations to warm arid ecosystems on six continents.

METHODS SUMMARY

Sequencing. Reads were collected from Applied Biosystems 3730xl, Roche 454 XLR and Illumina Genome Analyzer Iix machines at the Joint Genome Institute

(http://www.jgi.doe.gov/sequencing/protocols/protos_production.html) or Hudson-Alpha Institute and Beckman Coulter Genomics (BAC end sequence), and USDA-ARS Mid-South Area Genomics Laboratory (*G. longicalyx*, *G. arboreum* and *G. hirsutum*).

Assembly. Assembly of 80,765,952 sequence reads used a modification of Arachne v.20071016, integrating linear (15× genome coverage) and paired (3.1× genome coverage) Roche 454 libraries corrected using 41.9 Gb Illumina sequence, with 1.54× paired-end Sanger sequences from two subclone, six fosmid and two BAC libraries. Cotton genetic and physical maps, and *Vitis vinifera* and *T. cacao* synteny were used to identify 51 joins across 64 scaffolds to form the 13 chromosomes (Supplementary Discussion, section 1). The remaining scaffolds were screened for contamination to produce a final assembly of 1,033 scaffolds (19,735 contigs) and 761.4 Mb. Sequences are in NCBI for *G. raimondii* (BioProject accession PRJNA171262), *G. longicalyx* (accession F1-1, SRA061660), *G. herbaceum* (accession A1-97, SRA061243) and *G. hirsutum* (cultivar Acala Maxxa, SRS375727) genomes; *G. hirsutum* (SRA061240) and *G. barbadense* (SRA061309) fibre transcriptomes; *G. hirsutum* (SRA061456) seed transcriptomes; and *G. hirsutum* microRNAs (SRA061415).

Annotation. PERTRAN software was used to construct transcript assemblies from ~1.1 billion pairs of *G. raimondii* paired-end Illumina RNA-seq reads, 250 million *G. raimondii* single end reads, and 150 million *G. hirsutum* single end reads. PASA³⁰ was used to build transcript assemblies from 454 and Sanger resources (Supplementary Table 2.3). Loci were determined by transcript assembly and/or EXONERATE alignments of *A. thaliana*, cacao, rice, soybean, grape and poplar peptides to repeat-soft-masked *G. raimondii* genome using RepeatMasker. Gene models were predicted by three homology-based predictors (Supplementary Discussion, section 2.2). Best-scoring gene predictions were improved by PASA, then filtered on the basis of peptide homology or expressed-sequence-tag evidence to remove Pfam transposable element domain models. ClustalW alignments of amino acid sequences (Fig. 3) were used to guide coding sequence alignments. Phylogenetic trees were constructed by bootstrap neighbour-joining with a Kimura 2-parameter model using ClustalW2, assessing internal nodes with 1,000 replicates.

Full Methods and any associated references are available in the online version of the paper.

Received 17 August; accepted 21 November 2012.

- Jiang, C., Wright, R. J., El-Zik, K. M. & Paterson, A. H. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc. Natl Acad. Sci. USA* **95**, 4419–4424 (1998).
- Wendel, J. F. New world tetraploid cottons contain old-world cytoplasm. *Proc. Natl Acad. Sci. USA* **86**, 4132–4136 (1989).
- Wang, X. *et al.* The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genet.* **43**, 1035–1139 (2011).
- Senchina, D. S. *et al.* Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. *Mol. Biol. Evol.* **20**, 633–643 (2003).
- Wang, K. *et al.* The draft genome of a diploid cotton *Gossypium raimondii*. *Nature Genet.* **44**, 1098–1103 (2012).
- Carvalho, M. R., Herrera, F. A., Jaramillo, C. A., Wing, S. L. & Callejas, R. Paleocene Malvaceae from northern South America and their biogeographical implications. *Am. J. Bot.* **98**, 1337–1355 (2011).
- Jaillon, O. *et al.* The French-Italian Public Consortium for Grapevine Genome Characterization. The grapevine genome sequence suggests ancestral hexaploidization in the major angiosperm phyla. *Nature* **449**, 463–467 (2007).
- Bowers, J. E., Chapman, B. A., Rong, J. & Paterson, A. H. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433–438 (2003).
- Muravenko, O. V. *et al.* Comparison of chromosome BrdU–Hoechst–Giemsa banding patterns of the A1 and (AD)2 genomes of cotton. *Genome* **41**, 616–625 (1998).
- Jiao, Y. *et al.* A genome triplication associated with early diversification of the core eudicots. *Genome Biol.* **13**, R3 (2012).
- Fawcett, J. A., Maere, S. & Van de Peer, Y. Plants with double genomes might have had a better chance to survive the Cretaceous–Tertiary extinction event. *Proc. Natl Acad. Sci. USA* **106**, 5737–5742 (2009).
- Stracke, R., Werber, M. & Weisshaar, B. The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* **4**, 447–456 (2001).
- Adkisson, P. L., Niles, G. A., Walker, J. K., Bird, L. S. & Scott, H. B. Controlling Cotton's insect pests: a new system. *Science* **216**, 19–22 (1982).
- Wang, X., Tang, H. & Paterson, A. H. Seventy million years of concerted evolution of a homoeologous chromosome pair, in parallel in major Poaceae lineages. *Plant Cell* **23**, 27–37 (2011).
- Haigler, C. H., Betancur, L., Stiff, M. R. & Tuttle, J. R. Cotton fiber: a powerful single-cell model for cell wall and cellulose research. *Front. Plant Sci.* **3**, 1–7 (2012).
- Doblin, M. S., Pettolino, F. & Bacic, A. Plant cell walls: the skeleton of the plant world. *Funct. Plant Biol.* **37**, 357–381 (2010).
- Baulcombe, D. RNA silencing in plants. *Nature* **431**, 356–363 (2004).
- Matzke, M. A. & Birchler, J. A. RNAi-mediated pathways in the nucleus. *Nature Rev. Genet.* **6**, 24–35 (2005).
- Brodersen, P. & Voinnet, O. The diversity of RNA silencing pathways in plants. *Trends Genet.* **22**, 268–280 (2006).
- Kim, H. J. & Triplett, B. A. Cotton fiber germin-like protein. I. Molecular cloning and gene expression. *Planta* **218**, 516–524 (2004).
- Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* **39**, e118 (2011).
- Rong, J. *et al.* Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. *Genetics* **176**, 2577–2588 (2007).
- Wang, G. L., Dong, J. M. & Paterson, A. H. The distribution of *Gossypium hirsutum* chromatin in *G. barbadense* germ plasma: molecular analysis of introgressive plant breeding. *Theor. Appl. Genet.* **91**, 1153–1161 (1995).
- Richly, E. & Leister, D. NUMTs in sequenced eukaryotic genomes. *Mol. Biol. Evol.* **21**, 1081–1084 (2004).
- Reinisch, A. J. *et al.* A detailed RFLP map of cotton (*Gossypium hirsutum* × *Gossypium barbadense*): chromosome organization and evolution in a disomic polyploid genome. *Genetics* **138**, 829–847 (1994).
- Flagel, L. E., Wendel, J. F. & Udall, J. A. Duplicate gene evolution, homoeologous recombination, and transcriptome characterization in allopolyploid cotton. *BMC Genomics* **13**, 302 (2012).
- Wendel, J. F., Schnabel, A. & Seelanan, T. An unusual ribosomal DNA sequence from *Gossypium gossypoides* reveals ancient, cryptic, intergenomic introgression. *Mol. Phylogenet. Evol.* **4**, 298–313 (1995).
- Zhao, X. P. *et al.* Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Res.* **8**, 479–492 (1998).
- Cronn, R., Small, R. L., Haselkorn, T. & Wendel, J. F. Cryptic repeated genomic recombination during speciation in *Gossypium gossypoides*. *Evolution* **57**, 2475–2489 (2003).
- Haas, B. J. *et al.* Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements The work, conducted by the US Department of Energy Joint Genome Institute, is supported by the Office of Science of the US Department of Energy under Contract no. DE-AC02-05CH11231. The authors appreciate financial support from the US National Science Foundation (DBI 98-72630 to A.H.P., J.F.W., A.R.G.; DBI 02-11700 to J.F.W., A.H.P., J.U., A.R.G.; DBI 02-08311, IIP-0917856; IIP-1127755 to A.H.P.; IOS 1025947 to C.H.H.), USDA (ARS-58-6402-7-241, 58-6402-1-644 and 58-6402-1-645 to D.G.P.; ARS 6402-21310-003-00 to B.E.S.; NRI 00-52100-9685 and 02-35301-12045 to A.H.P.), Bayer CropScience and The Consortium for Plant Biotechnology Research (A.H.P.), Cotton, Inc. (P.W.C., D.C.J., A.H.P., D.M.S., A.V.-D., J.F.W.), Georgia State Support Committee (P.W.C., A.H.P.), Texas State Support Committee (R.J.W.), Pakistan–US Science and Technology Cooperation Program (P.W.C., S.M., A.H.P., M.u.R.), US–Egypt Science and Technology Cooperation Program (A.H.P., E.A.Z.), Fulbright Scholar Program (S.M., E.A.Z.), Conselho Nacional de Desenvolvimento Científico e Tecnológico PDJ150690/2012-6 (E.R.), Fundação de Amparo a Pesquisa Pensar Rio E-26/110.324/2010 (M.F.S.V.), Texas AgLife (D.M.S.), and Brigham Young University (BYU) Mentored Environment Grants (J.U.). RNA-seq reads were mapped by students on Marylou at the Fulton Supercomputer Center at BYU. We thank L. S. Dure III, G. O. Myers, J. McD Stewart, T. A. Wilkins and J. Zhu for co-endorsing the sequencing of *G. raimondii* by the US Department of Energy.

Author Contributions A.H.P., D.S.R., J.S. and D.G.P. conceived the study. J.S., J.G., D.S.R., K.C.S., S.D., M.V.D., C.L., L.N.R., B.E.S. and J.A.S. performed sequencing and associated clone manipulations. A.H.P., J.F.W., D.L., E.S.D., J.U., E.R., Z.W., A.H., L.V.H., R.H., D.M.S., A.V.-D. and T.Z. contributed unpublished data. A.H.P., J.F.W., H.Guo, H.Gundlach, J.J., D.J., D.L., S.S., J.U., M.-j.Y., R.B., W.C., A.D.-F., L.G., C.G., K.G., G.H., T.-h.L., J.L., L.L., T.L., B.S.M., J.T.P., A.W.R., E.R., E.S., X.T., C.X., J.W., Z.W., D.Z., L.Z., F.B., C.H.H., D.H., L.V.H., R.H., S.M., M.F.S.V., S.A.W., T.Z., E.S.D., D.S.R., X.W. and J.S. analysed data. A.H.P., J.F.W., D.L., E.S.D., K.F.X.M., D.G.P. and J.S. wrote the manuscript. All authors discussed results and commented on the manuscript.

Author Information Sequences have been deposited in NCBI for *G. raimondii* (BioProject accession PRJNA171262), *G. longicalyx* (accession F1-1, SRA061660), *G. herbaceum* (accession A1-97, SRA061243) and *G. hirsutum* (cultivar Acala Maxxa, SRS375727) genomes; *G. hirsutum* (SRA061240) and *G. barbadense* (SRA061309) fibre transcriptomes; *G. hirsutum* (SRA061456) seed transcriptomes; and *G. hirsutum* microRNAs (SRA061415). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.H.P. (paterson@plantbio.uga.edu) or J.S. (jschmutz@hudsonalpha.org).



This work is licensed under a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported licence. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-sa/3.0>

METHODS

Sequencing. Reads were collected with standard protocols (http://www.jgi.doe.gov/sequencing/protocols/protos_production.html) on Applied Biosystems 3730xl, Roche 454 XLR and Illumina Genome Analyzer (GA)IIx machines at the US Department of Energy Joint Genome Institute. Linear 454 data included standard XLR (47 runs, 16.868 Gb) and pre-release FLX+ data (5 runs, 3.262 Gb). Eight paired 454, 3–4-kilobase (kb) average insert size and one paired 12-kb average insert size were sequenced on standard XLR (23 runs, 5.931 Gb). One standard 400-base pair (bp) fragment library was sequenced at 2×150 (7 channels, 41.9 Gb) on an Illumina GAIIx. One 2.5-kb average insert size (405,024 reads, 286.1 Mb), one 6.5-kb average insert size library (374,125 reads, 263.0 Mb), six fosmid libraries (1,222,643 reads, 702.1 Mb) of 34–39-kb insert size, and two BAC libraries (107,520 reads, 77.5 Mb) of 98-kb and 115-kb (73,728 reads, 48.8 Mb) average insert size were sequenced on both ends for a total of 2,183,240 Sanger reads of 1.38 Gb of high-quality bases. FLX+ data were collected at the Roche Service Center. BAC end sequence (BES) was collected using standard protocols at the HudsonAlpha Institute.

Genome assembly and construction of pseudomolecule chromosomes. Organellar reads were removed by screening against mitochondria, chloroplast and ribosomal DNA. Any Roche 454 linear read <200 bp was discarded. Roche 454 paired reads in which either was shorter than 50 bp were discarded. An additional de-duplication step was applied to the 454 paired libraries that identifies and retains only one copy of each PCR duplicate. All remaining 454 reads were compared against a full Illumina GA2x run and any insertion/deletions in the 454 reads were corrected to match the Illumina alignments. The sequence reads were assembled using our modified version of Arachne v.20071016 (ref. 31) with parameters `maxcliq1 = 100`, `correct1_passes = 0` and `BINGE_AND_PURGE = True`, `bless = False`, `maxcliq1 = 200`, `BINGE_AND_PURGE = True`, `lap_ratio = 0.8`, `max_bad_look = 1000` (note Arachne error correction is on). This produced 1,263 scaffold sequences, with a scaffold L50 of 25.8 Mb, 58 scaffolds larger than 100 kb, and total genome size of 761.8 Mb. Scaffolds were screened against bacterial proteins, organelle sequences and non-redundant GenBank and removed if found to be a contaminant. Additional scaffolds were removed if they: (1) consisted of >95% 24-nucleotide sequences that occurred four other times in scaffolds larger than 50 kb; (2) contained only unanchored RNA sequences; or (3) were <1 kb in length.

The combination of BES/markers hybridized to fingerprint contigs³², 2,800 markers in a genetic map for the D genome in an *A₁D₁* plant³³ and 262 markers from the tetraploid genetic map³⁴, along with *Vitis vinifera* and *T. cacao* synteny was used to identify breaks in the initial assembly. Markers were aligned to the assembly using BLAT³⁵ (parameters: `-t = dna -q = dna -minScore = 200 -extendThroughN`). BES, physical map contigs, *V. vinifera* and *T. cacao* genes were aligned to the genome using BLAT³⁶. Scaffolds were broken if they contained linkage group/syntenic discontinuity coincident with an area of low BAC/fosmid coverage. A total of 13 breaks were executed, and 64 of the broken scaffolds were oriented, ordered and joined using 51 joins to form the final assembly containing 13 pseudomolecule chromosomes. Each chromosome join is padded with 10,000 missing nucleotides. The final assembly contains 1,033 scaffolds (19,735 contigs) that cover 761.4 Mb of the genome with a contig L50 of 135.6 kb and a scaffold L50 of 62.2 Mb.

The assembly size is near the centre of genome-size estimates of 880 Mb from flow cytometry³⁷, 630 Mb from Feulgen cytophotometry³⁸, and 650 Mb³⁹ and 770 Mb⁴⁰ from re-naturation kinetics.

Completeness of the euchromatic portion of the genome assembly was assessed using 65,506 *G. raimondii* complementary DNAs obtained from GenBank, which were aligned to the assembly using BLAT³ (parameters: `-t = dna -q = rna -extendThroughN`). The aim of the completeness analysis was to obtain a measure of completeness of the assembly, rather than a comprehensive examination of gene space. cDNAs were aligned to the assembly using BLAT³⁵ (parameters: `-t = dna -q = rna -extendThroughN`) and alignments that comprised $\geq 90\%$ base-pair identity and $\geq 85\%$ EST coverage were retained. The screened alignments indicate that 57,170 out of 63,506 (90.3%) cDNAs aligned to the assembly. The cDNAs that failed to align were primarily composed of stretches of polynucleotide

sequences that failed to generate non-random alignments to any plant or other organism in the NCBI as of the release date.

Annotation. A total of 85,746 transcript assemblies were constructed from about 1.1 billion pairs of D5 paired-end Illumina RNA-seq reads, 55,294 transcript assemblies from 250 million D5 single-end Illumina RNA-seq reads and 62,526 transcript assemblies from 150 million *G. hirsutum* cotton single-end Illumina RNA-seq reads. All these transcript assemblies were constructed using PERTRAN software (in preparation). In total, 120,929 transcript assemblies were built using PASA³⁰ from 56,638 D5 Sanger ESTs, 2.5 million D5 Roche 454 RNA-seq reads and all of the RNA-seq transcript assemblies. An additional 133,073 transcript assemblies were constructed using PASA from 296,214 *G. hirsutum* cotton Sanger ESTs and about 2.9 million *G. hirsutum* cotton 454 reads. The larger number of transcript assemblies from fewer *G. hirsutum* sequences is due to the fragmented nature of the assemblies. Loci were determined by transcript assembly alignments and/or EXONERATE alignments of peptides from *A. thaliana*, cacao, rice, soybean, grape and poplar peptides to repeat-soft-masked D5 genome using RepeatMasker (<http://www.repeatmasker.org>) with up to 2,000-bp extensions on both ends, unless extending into another locus on the same strand. Gene models were predicted by homology-based predictors, FGENESH⁴¹, FGENESH_EST (similar to FGENESH+, EST as splice site and intron input instead of peptide/translated open-reading frames) and GenomeScan⁴². The best scored predictions for each locus are selected using multiple positive factors including EST and peptide support, and one negative factor: overlap with repeats. The selected gene predictions were improved by PASA. Improvement includes adding untranslated regions, splicing correction, and adding alternative transcripts. PASA-improved gene model peptides were subject to peptide-homology analysis to above-mentioned proteomes in order to obtain Cscore and peptide coverage. Cscore is a peptide BLASTP score ratio mutual best hit BLASTP score and peptide coverage is highest percentage of peptide aligned to the best of homologues. PASA-improved transcripts were selected on the basis of Cscore, peptide coverage, EST coverage and its coding sequence (CDS) overlapping with repeats. The transcripts were selected if their Cscore was larger than or equal to 0.5 and peptide coverage larger than or equal to 0.5, or if it had EST coverage, but its CDS overlapping with repeats was less than 20%. For gene models whose CDS overlaps with repeats for more than 20%, its Cscore needed to be at least 0.9 and homology coverage at least 70% to be selected. The selected gene models were subject to Pfam analysis and gene models whose peptide was more than 30% in Pfam transposable element domains were removed. The final gene set had 37,505 protein-coding genes and 77,267 protein-coding transcripts.

- Jaffe, D. B. *et al.* Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res.* **13**, 91–96 (2003).
- Lin, L. *et al.* A draft physical map of a D-genome cotton species (*Gossypium raimondii*). *BMC Genomics* **11**, 389–417 (2010).
- Rong, J.-K. *et al.* A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). *Genetics* **166**, 389–417 (2004).
- Rong, J. *et al.* Comparative genomics of *Gossypium* and *Arabidopsis*: unraveling the consequences of both ancient and recent polyploidy. *Genome Res.* **15**, 1198–1210 (2005).
- Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **4**, 656–664 (2002).
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
- Hendrix, B. & Stewart, J. M. Estimation of the nuclear DNA content of *Gossypium* species. *Ann. Bot.* **95**, 789–797 (2005).
- Kadir, Z. B. A. DNA evolution in the genus *Gossypium*. *Chromosoma* **56**, 85–94 (1976).
- Geever, R., Katterman, F. & Endrizzi, J. DNA hybridization analyses of *Gossypium* allotetraploid and two closely related diploid species. *Theor. Appl. Genet.* **77**, 553–559 (1989).
- Walbot, V. & Dure, L. S. Developmental biochemistry of cotton seed embryogenesis and germination. VII. Characterization of cotton genome. *J. Mol. Biol.* **101**, 503–536 (1976).
- Salamov, A. A. & Solovyev, V. V. *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res.* **10**, 516–522 (2000).
- Yeh, R.-F., Lim, L. P. & Burge, C. Computational inference of homologous gene structures in the human genome. *Genome Res.* **11**, 803–816 (2001).

A prefrontal cortex–brainstem neuronal projection that controls response to behavioural challenge

Melissa R. Warden¹, Aslihan Selimbeyoglu^{1,2}, Julie J. Mirzabekov¹, Maisie Lo³, Kimberly R. Thompson¹, Sung-Yon Kim^{1,2}, Avishek Adhikari¹, Kay M. Tye^{1,4}, Loren M. Frank^{5,6} & Karl Deisseroth^{1,2,7,8,9}

The prefrontal cortex (PFC) is thought to participate in high-level control of the generation of behaviours (including the decision to execute actions¹); indeed, imaging and lesion studies in human beings have revealed that PFC dysfunction can lead to either impulsive states with increased tendency to initiate action², or to amotivational states characterized by symptoms such as reduced activity, hopelessness and depressed mood³. Considering the opposite valence of these two phenotypes as well as the broad complexity of other tasks attributed to PFC, we sought to elucidate the PFC circuitry that favours effortful behavioural responses to challenging situations. Here we develop and use a quantitative method for the continuous assessment and control of active response to a behavioural challenge, synchronized with single-unit electrophysiology and optogenetics in freely moving rats. In recording from the medial PFC (mPFC), we observed that many neurons were not simply movement-related in their spike-firing patterns but instead were selectively modulated from moment to moment, according to the animal's decision to act in a challenging situation. Surprisingly, we next found that direct activation of principal neurons in the mPFC had no detectable causal effect on this behaviour. We tested whether this behaviour could be causally mediated by only a subclass of mPFC cells defined by specific downstream wiring. Indeed, by leveraging optogenetic projection-targeting to control cells with specific efferent wiring patterns, we found that selective activation of those mPFC cells projecting to the brainstem dorsal raphe nucleus (DRN), a serotonergic nucleus implicated in major depressive disorder⁴, induced a profound, rapid and reversible effect on selection of the active behavioural state. These results may be of importance in understanding the neural circuitry underlying normal and pathological patterns of action selection and motivation in behaviour.

To execute an action that expends energy and requires vigorous effort under challenging conditions represents a consequential decision for an organism, particularly as such an action may not always represent the most adaptive behaviour. When a vigorous action pattern is selected despite extremely difficult circumstances (rather than a more energy-conserving passive or depressive-type pattern), an assessment may have occurred that the anticipated outcomes of the vigorous action would justify the expenditure of energy. Conversely, when an organism selects inactive behavioural patterns during challenging situations, the decision may represent anticipation that effort would probably be fruitless. Such anticipation leading to inaction can become maladaptive in human beings, with clinical symptoms including psychomotor retardation and hopelessness (core defining features of major depression, a disease with lifetime prevalence of nearly 20% and extensive socioeconomic ramifications⁵).

We sought to probe these high-level processes governing behavioural state selection with targeted control of restricted sets of circuit

elements in freely moving mammals. Increasing evidence suggests that the prefrontal cortex (PFC) could be involved in these behaviours; the PFC is responsible for coordinating thought and action, and has been shown to be critical for goal-oriented behaviour, planning and cognitive control^{6,7}—all of which are impaired in pathological states such as depression^{8–11}. Moreover, deep brain stimulation of the subcallosal cingulate region elicits antidepressant effects in treatment-resistant patients¹². Electrical stimulation of the rodent mPFC induces an antidepressant-like reduction in immobility in the forced swim test (FST)¹³, optogenetic stimulation of mixed excitatory and inhibitory neural populations in mPFC has an antidepressant-like effect in social defeat¹⁴, and mPFC in rodents has a role in mediating resilience¹⁵. Finally, neuroimaging studies in human patients have been instrumental in focusing attention on brain regions including PFC that exhibit abnormal activity in depression and melancholic states^{3,16,17}.

Despite these pioneering efforts exploring the role of the PFC, it is unclear which specific neural pathways are involved in real-time selection of effortful behavioural responses to challenging situations. The FST is relevant to this issue, as a widely employed behavioural test of depression-related responses in rodents¹⁸. In the FST, rodents are placed in a tank of water from which it is not possible to escape, and animals exhibit periods (epochs) of floating, which are thought to reflect states of passive coping or behavioural despair¹⁸, interspersed with epochs of active escape behaviour. Immobility in the FST is influenced by antidepressant drugs¹⁹ and stress²⁰. Transitions between active-escape and behavioural-despair states in the FST are clearly demarcated, in principle providing an unambiguous, instantaneous classification of a specific motivated behavioural state and an opportunity to investigate the neural dynamics underlying the decision to adopt an active behavioural response to challenge. However, to our knowledge, neural activity has never been recorded in behaving animals during the FST because of the technical obstacles of recording and controlling neural activity in a freely swimming animal. To address this challenge, we developed new methods for recording millisecond-precision neural and behavioural data alongside optogenetic control during the FST (Fig. 1).

We designed a magnetic induction method to detect individual swim kicks, in which the FST tank of water was surrounded by an induction coil and a small magnet was attached to the hind paw of the subject animal (Fig. 1a). During the FST, each kick induced a current in the coil (Fig. 1b); it was possible to cleanly isolate single kicks (Supplementary Fig. 1), and both kick frequency and automatically scored immobility corresponded well to manually scored immobility (Fig. 1c, d). In addition, we used this method to record mobile and immobile states during cage activity (Supplementary Fig. 2). To record well-isolated single units during swimming, tetrode microdrives or fixed wire arrays were waterproofed (Supplementary Methods). Under these conditions, we were reliably able to isolate single units during the FST (Fig. 1e); indeed,

¹Department of Bioengineering, Stanford University, Stanford, California 94305, USA. ²Neurosciences Program, Stanford University, Stanford, California 94305, USA. ³Bio-X Program, Stanford University, Stanford, California 94305, USA. ⁴Picower Institute for Learning & Memory, Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ⁵Department of Physiology, University of California San Francisco, San Francisco, California 94143, USA. ⁶W.M. Keck Center for Integrative Neuroscience, University of California San Francisco, San Francisco, California 94143, USA. ⁷Department of Psychiatry & Behavioral Sciences, Stanford University, Stanford, California 94305, USA. ⁸CNC Program, Stanford University, Stanford, California 94305, USA. ⁹Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, USA.

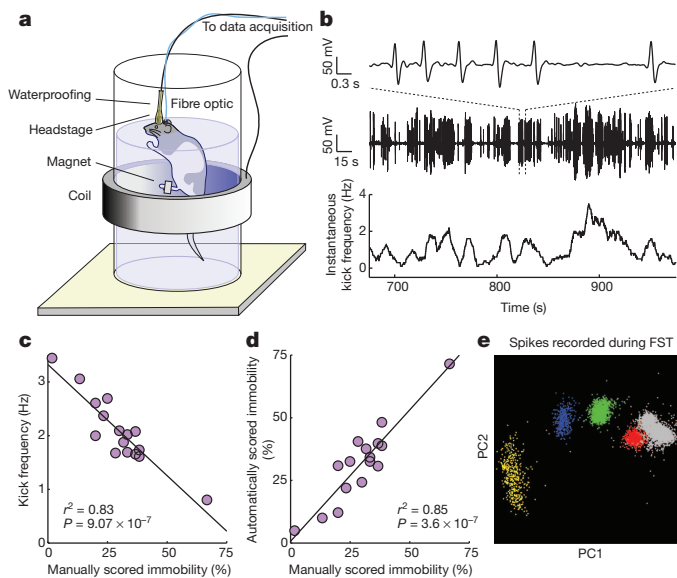


Figure 1 | The automated FST provides a high-temporal-resolution behavioural read-out that can be synchronized with simultaneously recorded neural data. **a**, A schematic of the automated FST. A coil of wire surrounds the tank of water and a magnet is attached to the rat's back paw. Movement of the magnet within the coil during swimming induces a current that can be recorded. To permit concurrent neural recordings, the headstage is waterproofed. An optical fibre can be included for simultaneous optical stimulation. **b**, Example FST coil voltage traces. Top, a 6-s coil trace showing individual kicks; middle, a 5-min coil trace; bottom, instantaneous kick frequency estimated from the 5-min coil trace. **c**, Average kick frequency corresponds well to manually scored immobility estimates. **d**, Estimates of FST immobility derived from the induction coil correspond tightly to manually scored immobility estimates. **e**, Four well-isolated single mPFC units recorded during the FST. PC, principal component (see Methods).

we were able to detect transitions between active escape behaviour and immobile states with high temporal precision, and to correlate these behaviours with ongoing neural activity (Fig. 2).

We recorded neural activity using either a 4-tetrode microdrive (6 rats) or a 24-electrode fixed-wire array (5 rats) targeted to the mPFC (Fig. 2a). Three epochs of data were recorded (Fig. 2b): a 15-min pre-FST epoch in a familiar cage, a 15-min epoch during the FST, and a 15-min post-FST epoch. We found that many mPFC neurons were strongly modulated during behaviour in a way that seemed to specifically reflect the decision to act or refrain from action during the FST. An example neuron is shown (Fig. 2c, d). This neuron was highly active during the mostly immobile pre- and post-FST epochs (98% and 94% immobile, respectively), but during the FST it stayed active during mobile states and was inhibited during immobile states. This neuron did not simply encode locomotor activity but was instead specifically inhibited during FST immobility corresponding to traditionally defined states of behavioural despair¹⁸.

Many neurons in the recorded population (23 out of 160, 14%; see Supplementary Methods) exhibited this surprising profile of activity. All rats exhibited minimal motor activity during the pre-FST epoch (greater than 88% immobility for all rats, average 97% immobility) and a moderate to high level of motor activity during the FST epoch (less than 79% immobility for all rats, average 39% immobility, Fig. 2e). Most recorded neurons (129 out of 160, 81%) showed a significant change in firing rate between pre-FST and FST epochs (Fig. 2f, top; statistics described in Fig. 2c and Methods). On average, this population of neurons was inhibited during the FST epoch (80 out of 129, 62%). Many neurons (70 out of 160, 44%) also showed a difference in firing rate between mobile and immobile states within the FST epoch (Fig. 2f, bottom). Most of these neurons were activated during mobile states and inhibited during immobile states (51 out of 70, 73%).

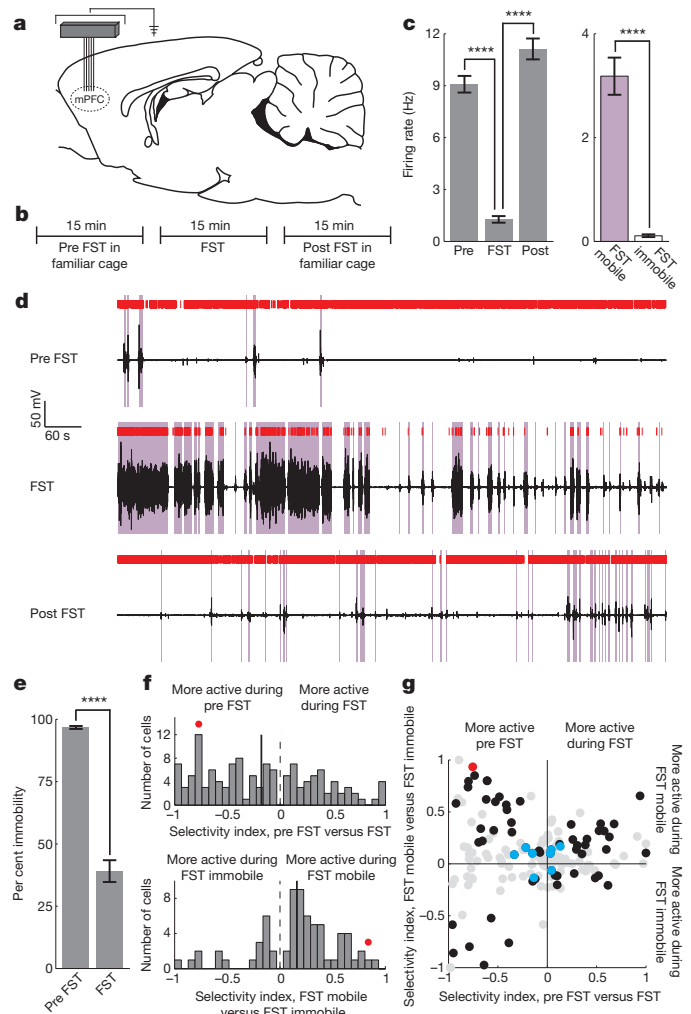


Figure 2 | Prefrontal neuronal activity encodes FST behavioural state. **a**, A tetrode microdrive or fixed wire array was implanted over the mPFC. **b**, Data were recorded pre FST, during the FST and post FST (for 15 min each). **c**, Bar plot of an example neuron that is inhibited during immobile states in the FST (Mann–Whitney *U*-test, $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$). **d**, Raster plot of the same neuron as in **c**. Coil voltage in black, mobile states in purple, spikes in red. Top, pre-FST activity; middle, activity during the FST; bottom, post-FST activity. **e**, Immobility during the pre-FST and FST test epochs (11 rats). **f**, Distribution of population selectivity indices (Supplementary Methods). Top, pre FST versus FST epochs. All neurons significantly selective for pre FST versus FST (as in panel **c**) are shown. Bottom, mobile versus immobile FST states. All neurons significantly selective for mobile versus immobile FST state are shown. **g**, Joint distribution of selectivity indices. Black circles, neurons selective for both task epoch and mobility; red circle, example neuron; blue circles, putative inhibitory fast-spiking neurons; grey circles, non-significantly selective neurons. All recorded neurons are shown. Error bars indicate s.e.m.

We then examined the joint distribution of epoch- and mobility-dependent neural selectivity among four quadrants (Fig. 2g), and found it to be highly asymmetric. The upper right and lower left quadrants exhibited a straightforward correspondence between motor activity and neural activity; for example, neurons in the upper right quadrant were more active during the largely mobile FST epoch than during the immobile pre-FST epoch, and, within the FST epoch, were more active during mobile states. The other two quadrants (the upper left and lower right quadrants) showed an inverted correspondence. In the upper left quadrant, neurons that were quieted during the more-active FST epoch were actually activated during escape behaviours within FST, but the opposite was true for neurons in the lower right quadrant. The profile of activity found within these groups was therefore not simply dependent on motor activity. We noted that putative

fast-spiking interneurons (Supplementary Methods) exhibited a reduced degree of modulation along both selectivity dimensions. Finally, we found that when mPFC neural activity was aligned to the onset of mobility epochs, firing on average preceded the onset of mobility (Supplementary Fig. 3).

The mPFC neurons that we recorded exhibited a range of selectivity profiles; therefore, it was not obvious that optogenetically activating local neurons in the mPFC would have a net effect on behaviour during the FST. To test this, we restricted opsin expression to Ca^{2+} /calmodulin-dependent protein kinase- α (CaMKII α)-expressing (chiefly excitatory) neurons within the mPFC using an adeno-associated viral vector (AAV5) expressing channelrhodopsin-2 fused to enhanced yellow fluorescent protein (ChR2-eYFP) under the control of the CaMKII α promoter. Virus was infused into the mPFC and fibre optics were implanted over the mPFC (Fig. 3a, b). We confirmed functional targeting of these neurons with optrode recordings in the mPFC under anaesthetic (Supplementary Fig. 4a), but surprisingly, when these neurons were illuminated in 2-min epochs during the FST (Supplementary Methods), we found that stimulation was not sufficient to cause even a slight reduction in FST immobility (Fig. 3c, Wilcoxon signed-rank test, $P = 0.23$) or a change in a control open-field-test (OFT) behaviour (Supplementary Fig. 4b). One interpretation of these results is that local mPFC neurons may correlate with, but are not causally involved in, the behavioural state

changes associated with effort-related mobility and immobility; alternatively, it could be that some local mPFC neurons are involved in this way, but others are not or are opposed in causal function, and when driven together no net effect on behaviour is seen. We therefore next considered that it could be possible to induce a change in this motivated behavioural state by restricting optogenetic stimulation to a reduced population of mPFC neurons.

The mPFC is known to project to several downstream brain regions that have been implicated in motivated behaviour and depression²¹; among these is the dorsal raphe nucleus (DRN)²², a serotonergic nucleus implicated in major depressive disorder⁴. The mPFC exerts control over both neural activity in the DRN and extracellular 5-hydroxytryptamine (5-HT; serotonin) levels^{15,23}, and antidepressant-like effects of mPFC electrical stimulation seem to depend on an intact 5-HT system¹³, but the projection from the mPFC to the DRN has not been directly shown to have an effect on behaviour. To specifically activate the mPFC–DRN projection, we first transduced excitatory neurons in the mPFC with AAV5 CaMKII α ::ChR2-eYFP (Fig. 3d), which led to robust ChR2-eYFP expression in mPFC axons in the DRN (Fig. 3e and Supplementary Fig. 5a). We restricted activation to the subpopulation of excitatory neurons in the mPFC that project to the DRN by implanting a fibre optic over the DRN and selectively illuminating this region (Fig. 3d). We confirmed functional

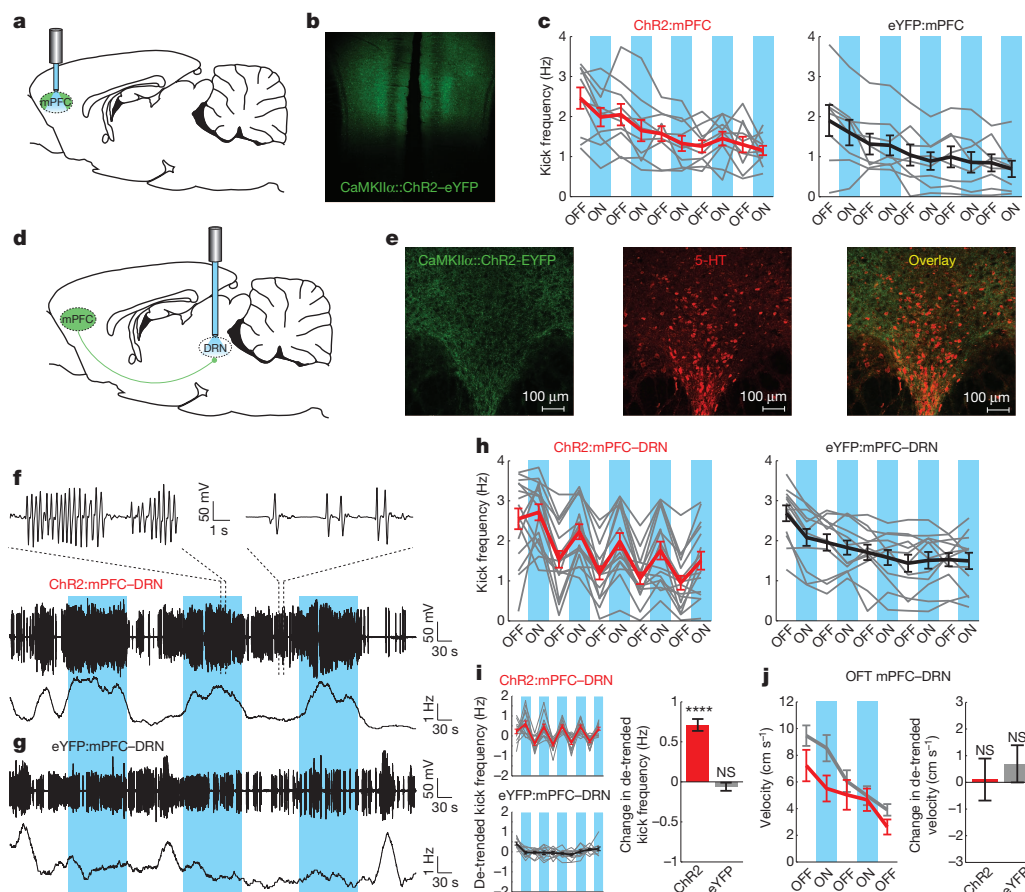


Figure 3 | Optogenetic stimulation of mPFC axons in the DRN, but not excitatory mPFC cell bodies, induces behavioural activation. **a**, ChR2-eYFP or eYFP-expressing mPFC principal neurons were illuminated directly. **b**, ChR2-eYFP fluorescence in the mPFC. **c**, FST kick frequency for the ChR2 in mPFC condition (ChR2:mPFC) (left, $n = 10$) and the eYFP in mPFC condition (eYFP:mPFC) (right, $n = 8$) rats. Grey lines, individual rats; thick lines, average for ChR2:mPFC (red) or eYFP:mPFC (black) rats; blue bars, light on. **d**, A fibre optic was implanted over the DRN after mPFC injection. **e**, ChR2-eYFP fluorescence in mPFC axons in the DRN (immunostained for 5-HT). **f**, FST behavioural data from one ChR2:mPFC-DRN rat. Top and middle, coil

voltage; bottom, kick frequency. **g**, FST data from one eYFP:mPFC-DRN rat. **h**, FST kick frequency for all rats. Left, ChR2:mPFC-DRN rats ($n = 16$); right, eYFP:mPFC-DRN rats ($n = 12$). **i**, Left, exponentially de-trended data from **h**; right, change in de-trended kick frequency from light-off to light-on epochs, ChR2:mPFC-DRN (red) and eYFP:mPFC-DRN (grey) rats. **j**, Left, velocity during stimulation in the open field test; right, change in de-trended velocity from light-off to light-on epochs. Red, ChR2:mPFC-DRN rats ($n = 12$); grey, eYFP:mPFC-DRN rats ($n = 12$). NS, not significant. Error bars indicate s.e.m.

targeting with optrode recordings under anaesthetic (Supplementary Fig. 5b).

When the axons of Chr2-eYFP-expressing mPFC neurons in the DRN were stimulated during the FST, a profound change in effortful behaviour resulted. Example induction-coil behavioural traces from two rats are shown (one Chr2-eYFP and one eYFP rat, Fig. 3f, g), demonstrating a robust increase in kick frequency during each light epoch in the Chr2-eYFP case. This behavioural effect was present in most Chr2-eYFP rats (Wilcoxon signed-rank test, $P = 1.04 \times 10^{-11}$) but not eYFP rats (Wilcoxon signed-rank test, $P = 0.39$), and was rapid, reversible and repeatable (Fig. 3h, i). Importantly, stimulation of this projection did not affect non-specific locomotor activity in the open field in either Chr2-eYFP rats (Wilcoxon signed-rank test, $P = 0.59$) or eYFP rats (Wilcoxon signed-rank test, $P = 0.71$, Fig. 3j). This result shows the importance of resolving neural subpopulations *in vivo* defined by anatomical presence of axonal projection target region, and illustrates a causal role of a specific mPFC-to-brainstem neural pathway in driving this motivated behavioural response to a challenging environment. Additional experiments addressing the effect of mPFC-DRN stimulation on mPFC neural activity are described in Supplementary Figs 6–9.

The mPFC projection to the DRN sends sparse collaterals to other brain regions²⁴; accordingly, we next blocked incoming glutamatergic synaptic activity in the DRN during stimulation of DRN-projecting mPFC axons (Supplementary Fig. 10). Glutamate receptor antagonists blocked stimulation-driven behavioural activation, consistent with the idea that activation of the mPFC-DRN synapse itself is necessary for the stimulation-induced increase in mobility. Additionally, inhibition of mPFC axons in the DRN led to a lasting decrease in steady-state mobility in the FST (Supplementary Fig. 11), pointing to the necessity of this pathway in normal behaviour.

We considered the possibility that specificity of the effect of mPFC on DRN would not be fully captured by general activation of the downstream region; we tested this question directly by transducing neurons in the DRN with Chr2-eYFP under the control of the human synapsin 1 promoter (AAV5 human synapsin promoter fragment (hSyn)::Chr2-eYFP), which transduced both 5-HT and GABA (γ -aminobutyric acid) neurons in the DRN (Supplementary Fig. 12), and implanted a fibre optic directly above the DRN (Fig. 4a–c). When Chr2-eYFP-expressing DRN cell bodies were illuminated directly, rats exhibited behavioural activation during the FST (Wilcoxon signed-rank test, $P = 5.63 \times 10^{-8}$), whereas eYFP-expressing rats did not (Wilcoxon signed-rank test, $P = 0.84$, Fig. 4d, e). However, direct activation of the DRN, unlike stimulation of the mPFC-DRN projection, led to a general increase in locomotor activity in the OFT in Chr2-eYFP rats (Wilcoxon signed-rank test, $P = 0.02$) but not eYFP rats (Wilcoxon signed-rank test, $P = 0.31$, Fig. 4f). It is likely that mPFC-DRN stimulation and direct DRN-cell-body stimulation activate different sub-networks within the DRN, which may explain these dissimilar behavioural results.

Finally, we targeted the projection from the mPFC to the lateral habenula²⁵, a region known to have an important role in motivated behaviour and depression^{26,27}, and found that activation of this specific projection had the opposite effect on escape-related behaviour in the FST. As above, we infused AAV5 CaMKII α ::Chr2-eYFP into the mPFC and implanted bilateral fibre optics over the lateral habenula (Fig. 4g–i). When Chr2-eYFP-expressing lateral-habenula-projecting mPFC axons were illuminated during the FST, rats showed a rapid and reversible decrease in mobility (Wilcoxon signed-rank test, $P = 3.63 \times 10^{-4}$), whereas control eYFP-expressing rats did not (Wilcoxon signed-rank test, $P = 0.29$, Fig. 4j–l, Supplementary Fig. 13).

Here, we have probed both neural correlates and causal neural pathways involved in the selection of an effortful motivated behavioural pattern during challenging circumstances, using novel technology that permits electrical recordings and optogenetic control in the FST in combination with high-speed reporting of behavioural state. We

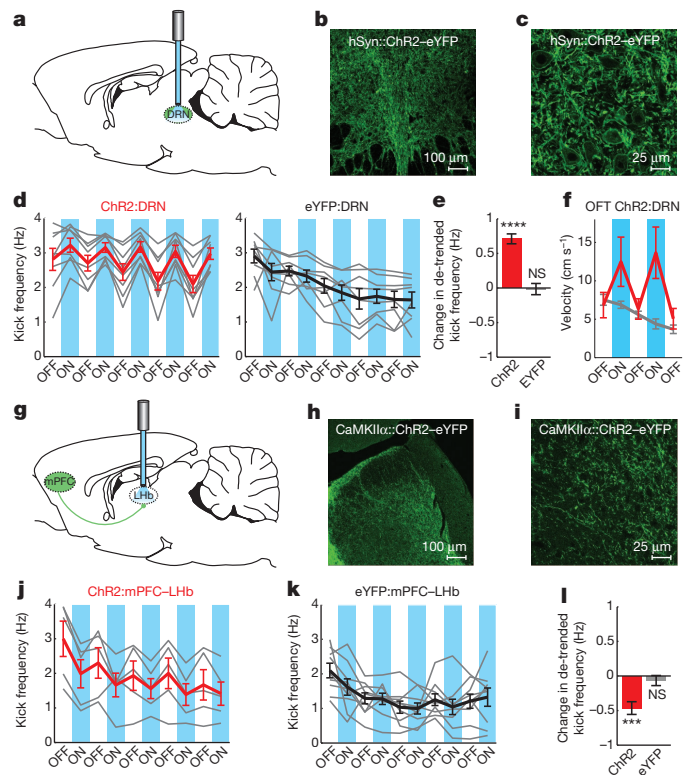


Figure 4 | Behavioural activation resulting from stimulation of DRN-projecting mPFC axons is specific to the mPFC-DRN synapse. **a**, A fibre optic was implanted over Chr2- or eYFP-expressing neurons in the DRN. **b**, Image ($\times 20$) of Chr2-eYFP-expressing DRN neuronal cell bodies. **c**, DRN image. **d**, FST kick frequency for all rats. Left, Chr2:DRN rats ($n = 8$); right, eYFP:DRN rats ($n = 8$); grey lines, individual rats; thick lines, average for Chr2:DRN (red) or eYFP:DRN (black) rats; blue bars, light on. **e**, De-trended change in kick frequency from light-off to light-on epochs, Chr2:DRN (red) and eYFP:DRN (grey) rats. **f**, Velocity during stimulation in the open field test. Red, Chr2:DRN ($n = 12$); grey, eYFP:DRN ($n = 12$) rats. **g**, Fibre optics were implanted bilaterally over the lateral habenula (LHb) to activate Chr2-expressing LHb-projecting mPFC axons. **h**, Chr2-eYFP-expressing mPFC axons in the LHb. **i**, LHb image. **j**, FST kick frequency for all rats. Left, Chr2:mPFC-LHb rats ($n = 5$); right, eYFP:mPFC-LHb rats ($n = 9$). **k**, De-trended change in kick frequency from light-off to light-on epochs in the FST. Error bars indicate s.e.m.

have shown the existence of different physiologically defined mPFC neural populations—one selectively inhibited during passive coping/behavioural despair-like states, and the other selectively activated during these states. We have also shown that, although general activation of CaMKII α -expressing neurons in the mPFC does not have a net effect on this behaviour in rats, selective activation in DRN or lateral habenula of projecting mPFC neurons elicits distinct, rapid and reversible effects on selection of the active behavioural state. These results describe the neural dynamics associated with the behavioural response to challenge and demonstrate the causal importance of mPFC control of downstream targets in implementing this response, with implications for understanding both normal and pathological states of decision pathway refinement, likely-outcome assessment, and behavioural pattern selection.

METHODS SUMMARY

Male Long-Evans rats were implanted with either a 4-tetrode microdrive or a 24-electrode fixed wire array targeted to the mPFC. Tetrodes were adjusted daily. Prior to the start of recordings rats were anaesthetized for 10 min to facilitate waterproofing of the headstage and electrodes, and were subsequently allowed to recover for at least 1 h. Data were analysed in Matlab and Neuroexplorer with custom-written software.

The pAAV-CaMKII α ::hChR2(H134R)-eYFP, pAAV-CaMKII α ::eYFP, pAAV-CaMKII α ::eNpHR3.0-eYFP and pAAV-hSyn::hChR2(H134R)-eYFP plasmids were designed and constructed by standard methods and packaged as AAV5. Virus was injected into the mPFC or the DRN. Maps and clones are available at <http://www.stanford.edu/group/dlab/optogenetics>.

The mPFC was virally transduced and an optical fibre was surgically implanted in separate surgeries over the DRN or the lateral habenula to allow for selective illumination of mPFC axons. Cannulae were used instead for pharmacology experiments. Virus was allowed to express for a minimum of 4 months after injection for projection-targeting experiments. Behavioural data were collected 7–10 days after fibre implantation.

To confirm opsin expression, coronal brain slices were prepared for immunohistochemistry and optical microscopy. Brain sections were stained for 4', 6-diamidino-2-phenylindole (DAPI) and either rabbit anti-5-HT or rabbit anti-GABA. Sections were imaged with a confocal microscope.

Received 18 January; accepted 24 September 2012.

Published online 18 November 2012.

- McGuire, J. T. & Botvinick, M. M. Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl Acad. Sci. USA* **107**, 7922–7926 (2010).
- Ridderinkhof, K. R., van den Wildenberg, W. P. M., Segalowitz, S. J. & Carter, C. S. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain Cogn.* **56**, 129–140 (2004).
- Mayberg, H. S. *et al.* Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am. J. Psychiatry* **156**, 675–682 (1999).
- Maes, M. & Meltzer, H. In *Psychopharmacology: the Fourth Generation of Progress* (eds Bloom, F. E. & Kupfer, D. J.) 933–944 (Raven Press, 1995).
- Kessler, R. C. *et al.* Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry* **6**, 168–176 (2007).
- Miller, E. K. & Cohen, J. D. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* **24**, 167–202 (2001).
- Fuster, J. M. *The Prefrontal Cortex, Fourth Edition* (Academic Press, 2008).
- Elliott, R. *et al.* Prefrontal dysfunction in depressed patients performing a complex planning task: a study using positron emission tomography. *Psychol. Med.* **27**, 931–942 (1997).
- Austin, M.-P. Cognitive deficits in depression: possible implications for functional neuropathology. *Br. J. Psychiatry* **178**, 200–206 (2001).
- Ingram, R. E., Bernet, C. Z. & McLaughlin, S. C. Attentional allocation processes in individuals at risk for depression. *Cognit. Ther. Res.* **18**, 317–332 (1994).
- Dalgleish, T. & Watts, F. N. Biases of attention and memory in disorders of anxiety and depression. *Clin. Psychol. Rev.* **10**, 589–604 (1990).
- Mayberg, H. S. *et al.* Deep brain stimulation for treatment-resistant depression. *Neuron* **45**, 651–660 (2005).
- Hamani, C. *et al.* Antidepressant-like effects of medial prefrontal cortex deep brain stimulation in rats. *Biol. Psychiatry* **67**, 117–124 (2010).
- Covington, H. E. *et al.* Antidepressant effect of optogenetic stimulation of the medial prefrontal cortex. *J. Neurosci.* **30**, 16082–16090 (2010).
- Amat, J. *et al.* Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nature Neurosci.* **8**, 365–371 (2005).
- Drevets, W. C. Neuroimaging and neuropathological studies of depression: implications for the cognitive-emotional features of mood disorders. *Curr. Opin. Neurobiol.* **11**, 240–249 (2001).
- Baxter, L. R. *et al.* Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Arch. Gen. Psychiatry* **46**, 243–250 (1989).
- Porsolt, R. D., Le Pichon, M. & Jalfre, M. Depression: a new animal model sensitive to antidepressant treatments. *Nature* **266**, 730–732 (1977).
- Cryan, J. F., Valentino, R. J. & Lucki, I. Assessing substrates underlying the behavioral effects of antidepressants using the modified rat forced swimming test. *Neurosci. Biobehav. Rev.* **29**, 547–569 (2005).
- Willner, P. Chronic mild stress (CMS) revisited: consistency and behavioural-neurobiological concordance in the effects of CMS. *Neuropsychobiology* **52**, 90–110 (2005).
- Vertes, R. P. Differential projections of the infralimbic and prelimbic cortex in the rat. *Synapse* **51**, 32–58 (2004).
- Gonçalves, L., Nogueira, M. I., Shammah-Lagnado, S. J. & Metzger, M. Prefrontal afferents to the dorsal raphe nucleus in the rat. *Brain Res. Bull.* **78**, 240–247 (2009).
- Celada, P., Puig, M. V., Casanovas, J. M., Guillazo, G. & Artigas, F. Control of dorsal raphe serotonergic neurons by the medial prefrontal cortex: involvement of serotonin-1A, GABA(A), and glutamate receptors. *J. Neurosci.* **21**, 9917–9929 (2001).
- Gabbott, P. L. A., Warner, T. A., Jays, P. R. L., Salway, P. & Busby, S. J. Prefrontal cortex in the rat: projections to subcortical autonomic, motor, and limbic centers. *J. Comp. Neurol.* **492**, 145–177 (2005).
- Kim, U. & Lee, T. Topography of descending projections from anterior insular and medial prefrontal regions to the lateral habenula of the epithalamus in the rat. *Eur. J. Neurosci.* **35**, 1253–1269 (2012).
- Matsumoto, M. & Hikosaka, O. Representation of negative motivational value in the primate lateral habenula. *Nature Neurosci.* **12**, 77–84 (2009).
- Sartorius, A. *et al.* Remission of major depression under deep brain stimulation of the lateral habenula in a therapy-refractory patient. *Biol. Psychiatry* **67**, e9–e11 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank H. Mayberg, R. Malenka, L. Gunaydin, J. Mattis, I. Ellwood and I. Witten for helpful comments on the manuscript; I. Ellwood, I. Witten, R. Airan, L. Meltzer, M. Roy, V. Gradinaru, A. Andalman, T. Davidson, R. Durand, M. Bower and M. Carr for useful discussions; and all members of the K.D. laboratory for their support. We are grateful to S. Pak, C. Ramakrishnan and C. Perry for technical assistance. This work was supported by the Wieggers Family Fund (K.D.), NARSAD (M.R.W. and K.R.T.), Stanford Graduate Fellowship (A.S.), Samsung Scholarship (S.-Y.K.), Berry Foundation Fellowship (A.A.), NIMH (1F32MH088010-01, K.M.T.), and NIMH, NIDA, the DARPA REPAIR Program, the Keck Foundation, the McKnight Foundation, the Yu, Snyder, Tarlton and Alice Woo Foundations, and the Gatsby Charitable Foundation (K.D.).

Author Contributions M.R.W., L.M.F. and K.D. contributed to study design with assistance from A.S. and K.M.T. M.R.W., L.M.F. and K.D. contributed to data interpretation and manuscript revision. M.R.W., A.S., K.M.T., J.J.M., M.L., K.R.T., S.-Y.K. and A.A. contributed to data collection. M.R.W. coordinated all experiments, developed the induction coil and forced swim test electrophysiology methods, and performed all behavioural and *in vivo* electrophysiology analyses. K.D. supervised all aspects of the project. M.R.W. and K.D. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.R.W. (mwarden@stanford.edu) or K.D. (deissero@stanford.edu).

Layered reward signalling through octopamine and dopamine in *Drosophila*

Christopher J. Burke^{1*}, Wolf Huetteroth^{2*}, David Oswald², Emmanuel Perisse², Michael J. Krashes^{1†}, Gaurav Das², Daryl Gohl³, Marion Silies³, Sarah Certel⁴ & Scott Waddell^{1,2}

Dopamine is synonymous with reward and motivation in mammals^{1,2}. However, only recently has dopamine been linked to motivated behaviour and rewarding reinforcement in fruitflies^{3,4}. Instead, octopamine has historically been considered to be the signal for reward in insects^{5–7}. Here we show, using temporal control of neural function in *Drosophila*, that only short-term appetitive memory is reinforced by octopamine. Moreover, octopamine-dependent memory formation requires signalling through dopamine neurons. Part of the octopamine signal requires the α -adrenergic-like OAMB receptor in an identified subset of mushroom-body-targeted dopamine neurons. Octopamine triggers an increase in intracellular calcium in these dopamine neurons, and their direct activation can substitute for sugar to form appetitive memory, even in flies lacking octopamine. Analysis of the β -adrenergic-like OCT β 2R receptor reveals that octopamine-dependent reinforcement also requires an interaction with dopamine neurons that control appetitive motivation. These data indicate that sweet taste engages a distributed octopamine signal that reinforces memory through discrete subsets of mushroom-body-targeted dopamine neurons. In addition, they reconcile previous findings with octopamine and dopamine and suggest that reinforcement systems in flies are more similar to mammals than previously thought.

Fruitfly octopamine is synthesized from tyrosine via two steps catalysed by tyrosine decarboxylase (TDC) and tyramine β -hydroxylase (Tbh)^{8,9}. The *Tdc2* gene encodes the neuronal TDC and a *Tdc2-GAL4* construct (where *GAL4* is cloned downstream of a *Tdc2* promoter fragment) can be used to label and manipulate many of the octopamine neurons⁸. Although *Tbh* mutant *Drosophila* that lack octopamine cannot form appetitive memory⁷, the precise role of octopamine release is currently unknown.

We tested whether octopamine neurons were required for appetitive olfactory conditioning with sucrose reinforcement by blocking them throughout the experiment using *Tdc2-GAL4*-driven *UAS-shibire^{ts1}* (*UAS-shi^{ts1}*, where the *shi^{ts1}* sequence is cloned downstream of the upstream activation sequence (*UAS*))¹⁰. The *UAS-shi^{ts1}* transgene allows temporary blockade of synaptic transmission from specific neurons by shifting flies from the permissive temperature <25 °C to the restrictive >29 °C. We assayed *Tdc2-GAL4;UAS-shi^{ts1}* flies in parallel with *GAL4* driver, *UAS-shi^{ts1}* transgene and wild-type flies for comparison (Fig. 1a). All flies were incubated at 31 °C to disrupt output from octopamine neurons for 30 min before being trained and tested for 3 h appetitive memory at 31 °C. Surprisingly, no defects were apparent.

Sweet taste and nutrient value both contribute to appetitive memory reinforcement in *Drosophila*^{11,12}. We reasoned that octopamine blockade might lack consequence if octopamine only represents sweet taste and nutrient value provides sufficient reinforcement. We therefore blocked *Tdc2* neurons while training flies with arabinose, a sweet

but non-nutritious sugar¹¹ (Fig. 1b). All flies were trained and tested for 3 min memory at 31 °C. In this case, memory of *Tdc2-GAL4;UAS-shi^{ts1}* flies was significantly impaired compared to all control groups. Importantly, no significant differences were apparent between groups trained and tested at 25 °C (Fig. 1c). To further challenge a nutrient bypass model we blocked octopamine neurons while flies were conditioned with arabinose supplemented with nutritious sorbitol¹¹. No differences were apparent (Fig. 1d), similar to blocking octopamine

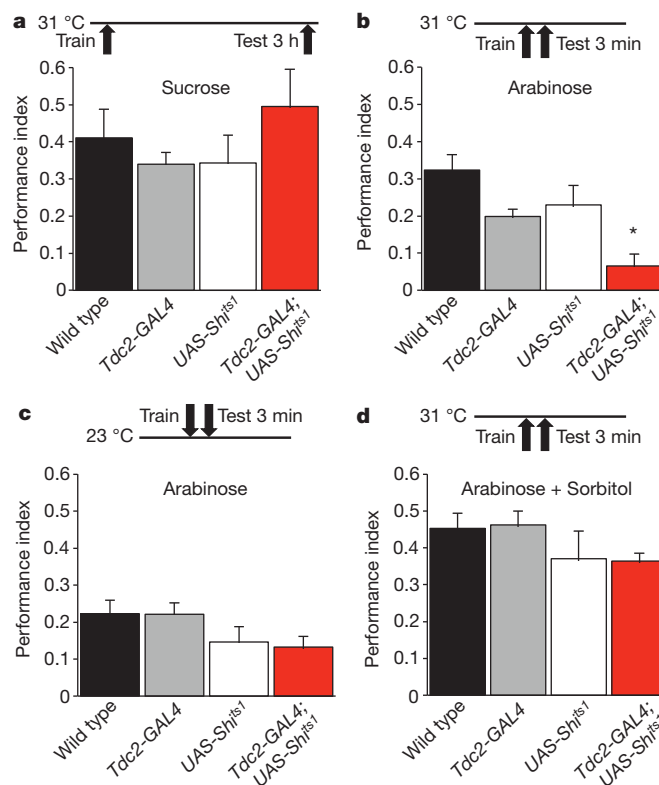


Figure 1 | Octopamine mediates the short-term reinforcing effects of sweet taste. **a**, Blocking octopamine neurons with *Tdc2-GAL4/UAS-shi^{ts1}* during conditioning with sucrose has no effect (all $P > 0.4$, $n \geq 6$). **b**, Blocking octopamine neurons during conditioning with arabinose significantly impairs appetitive learning ($P < 0.05$, $n \geq 10$). **c**, No significant defect is observed when flies are conditioned with arabinose at 23 °C (all $P > 0.1$, $n \geq 8$). **d**, Blocking octopamine neurons during conditioning with sorbitol-supplemented arabinose has no significant effect (all $P > 0.4$, $n = 8$). Data are mean \pm standard error of the mean (s.e.m.). Asterisks denote significant difference between marked groups and all others (all $P < 0.05$, ANOVA).

¹Department of Neurobiology, University of Massachusetts Medical School, 364 Plantation Street, Worcester, Massachusetts 01605, USA. ²Centre for Neural Circuits and Behaviour, The University of Oxford, Tinsley Building, Mansfield Road, Oxford OX1 3SR, UK. ³Department of Neurobiology, Stanford University, Stanford, California 94305, USA. ⁴Division of Biological Sciences and Center for Structural and Functional Neuroscience, University of Montana, Missoula, Montana 59812, USA. [†]Present address: Beth Israel Deaconess Medical Center, Harvard Medical School, Center for Life Sciences, 3 Blackfan Circle, Boston, Massachusetts 02215, USA.

*These authors contributed equally to this work.

neurons in flies conditioned with sweet and nutritious sucrose (Fig. 1a). These data are consistent with octopamine only conveying the reinforcing effects of sweet taste and with nutrient value being sufficient for appetitive learning¹².

To determine whether octopamine provides instructive reinforcement we conditioned flies with odour presentation paired with artificial octopamine neuron activation, achieved by expressing *UAS-dTrpA1* with *Tdc2-GAL4*. *dTrpA1* (also known as *TrpA1*) encodes a transient receptor potential (TRP) channel that conducts Ca^{2+} and depolarizes neurons when flies are exposed to temperature $>25^\circ\text{C}$ ¹³. *Ad-libitum*-fed wild-type, *Tdc2-GAL4*, *UAS-dTrpA1* and *Tdc2-GAL4;UAS-dTrpA1* flies were conditioned by presenting an odour with activating 31°C , and immediately tested for memory (Fig. 2a). *Tdc2-GAL4;UAS-dTrpA1* flies exhibited robust appetitive memory that was statistically different from all other groups (Fig. 2b). Significant memory remained at 30 min (Fig. 2c) in satiated flies but was statistically indistinguishable from all other groups at 3 h, even in hungry flies (Fig. 2d). Therefore appetitive memory implanted with octopamine neuron activation is short-lived. *Tdc2-GAL4* is expressed in neurons that contain and could release tyramine, either alone or together with octopamine (Supplementary Fig. 1). To confirm that artificial learning requires octopamine we stimulated *Tdc2* neurons in *Tbh* mutant flies⁹ that cannot synthesize octopamine from tyramine (Fig. 2e). No learning was observed, indicating that octopamine release is required for artificial learning.

Although octopamine neuron innervation of the mushroom body (MB) is relatively sparse in the γ lobe, heel and calyx¹⁴ (Supplementary Fig. 1), previous work suggests that MB neurons are probably the eventual destination of appetitive reinforcement signals^{4,7,15,16}. We therefore used the *NP7088*, *0665* and *0891-GAL4* lines to investigate the role of the four individual classes of octopamine (OA) neurons that innervate the MB¹⁴: OA-VUMa2, OA-VPM3, OA-VPM4 and OA-VPM5 (Fig. 2f). *NP7088* neurons visualized using *UAS-mCD8::GFP* (that is, expressing the green fluorescent protein (GFP)) broadly overlap with *Tdc2-GAL4* neurons in the brain (Fig. 2g), but *NP7088-GAL4* does not label OA-VPM5 neurons¹⁴. *0665-GAL4* (ref. 17) driven GFP expression is even more restricted and labels the OA-VPM3 and OA-VPM4 neurons (Fig. 2h). Finally, *0891-GAL4* (ref. 17) only labels OA-VPM4 (Fig. 2i).

Activating these subpopulations of MB-innervating octopamine neurons during odour presentation did not form appetitive memory (Fig. 2j). Similarly, blocking them (*NP7088*, *0665* or *0891-GAL4*) using *UAS-shi^{ts1}* did not significantly impair arabinose-reinforced memory (Supplementary Fig. 2). Importantly, these data indicate that the fly equivalent of the bee VUMnx1 neuron⁵, OA-VUMa2 (ref. 14), and the other MB-innervating neurons covered by these drivers, are neither sufficient nor essential for conditioned olfactory approach behaviour in flies. Instead, the data indicate that either the calyx-innervating OA-VPM5 neurons are critical, or a more distributed octopamine signal involving other non-MB-innervating octopamine neurons is required for appetitive reinforcement.

One study implicated the DopR dopamine receptor in appetitive memory. Flies with the *dumb¹* mutation have impaired appetitive memory that can be restored by expressing DopR in the MB¹⁶. We therefore tested whether memory formation with octopamine neuron activation required DopR (Fig. 3a). No significant memory was observed in any group carrying *dumb¹*. Therefore a functional dopamine system is required to form appetitive memory with octopamine, suggesting that dopamine is downstream of octopamine in appetitive memory processes.

A recent study implicated dopamine neurons in the PAM (paired anterior medial) cluster in appetitive reinforcement⁴. We independently identified the *0273-GAL4* and *0104-GAL4* lines in the InSITE collection¹⁷ that drive *UAS-mCD8::GFP* in subsets of PAM dopamine neurons that innervate the MB (Fig. 3b–e and Supplementary Fig. 3a). Co-labelling brains with *UAS-mCD8::GFP* and anti-tyrosine hydroxylase

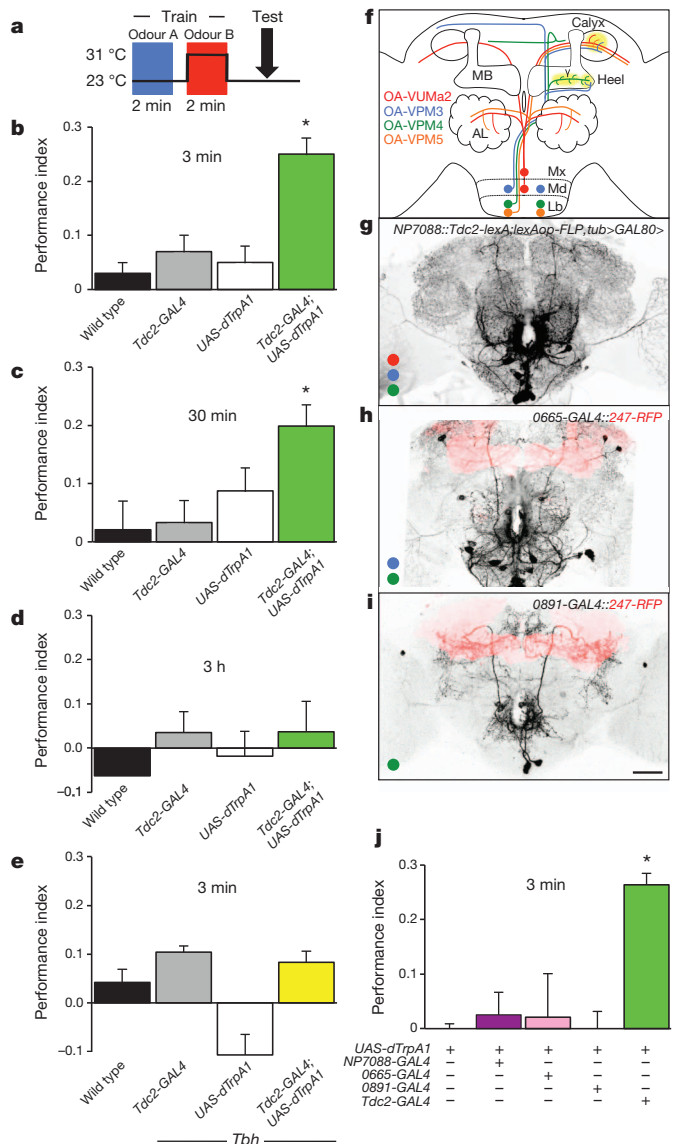


Figure 2 | Octopamine neuron stimulation can replace sugar presentation during conditioning to form short-term appetitive memory.

a, Conditioning protocol pairing a 2 min odour presentation with heat-activation (red) of *UAS-dTrpA1* expressing neurons. **b**, *Tdc2-GAL4/UAS-dTrpA1* driven octopamine neuron activation contingent with odour presentation forms appetitive olfactory memory in satiated flies ($P < 0.001$, $n \geq 14$). **c**, Implanted memory remains significant 30 min after training in satiated flies ($P < 0.05$, $n \geq 8$). **d**, No memory is observed at 3 h, even in hungry flies ($P > 0.5$, $n = 6$). **e**, Implanting memory with *Tdc2-GAL4* neuron stimulation requires octopamine. Artificial conditioning does not form significant memory in hungry *Tbh* flies ($P > 0.05$, $n \geq 8$). **f**, Schematic of all four octopamine neurons that innervate the MB calyx (OA-VUMa2, OA-VPM5, plus the antennal lobe, AL), heel (OA-VPM4, plus the MB γ lobe), or calyx and heel (OA-VPM3). Somata reside in the maxillary (Mx), mandibular (Md), or labial (Lb) neuromere. **g**, *NP7088-GAL4* expresses in many *Tdc2*-positive octopamine neurons. Projection of octopamine neurons common to *Tdc2-GAL4/Tdc2-lexA* and *NP7088-GAL4* revealed by genetic intersection. MB-octopamine cell types in each *GAL4* are colour-coded. **h**, *0665-GAL4* labels MB-innervating OA-VPM3 and OA-VPM4 neurons; 247-RFP labelled MB (red). **i**, *0891-GAL4* specifically labels MB-innervating OA-VPM4 neurons. Scale bar, 50 μm . **j**, Stimulating octopamine neuron subsets cannot replace sugar presentation in appetitive conditioning. ($P > 0.05$, $n \geq 6$).

(TH) antibody revealed that *0273-GAL4* is expressed in all of the ~ 130 dopamine neurons in the PAM cluster (Fig. 3c), whereas *0104-GAL4* labels a subset of 40 PAM dopamine neurons (Fig. 3e). Importantly,

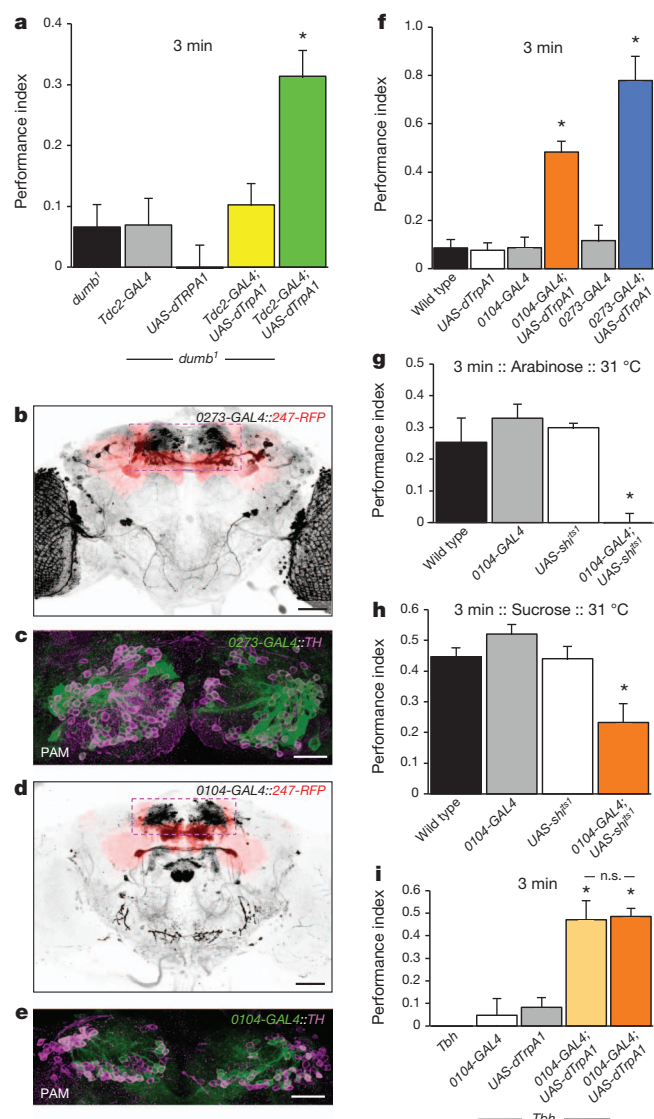


Figure 3 | Reinforcing dopamine neurons are functionally downstream of octopamine-dependent reinforcement. **a**, Memory cannot be implanted with *Tdc2* neuron stimulation in *dumb¹* (*DopR*) flies (all $P > 0.05$ except *Tdc2-GAL4/UAS-dTrpA1* control, $P < 0.001$, $n \geq 8$). **b**, 0273-GAL4 labels PAM dopamine neurons (dashed box) that innervate the MB (red). **c**, 0273-GAL4 labels all ~130 TH-positive PAM dopamine neurons. **d**, 0104-GAL4 labels PAM dopamine neurons (dashed box). **e**, 0104-GAL4 labels ~40 TH-positive PAM dopamine neurons. Scale bar, 50 μm (**b** and **d**), 20 μm (**c** and **e**). **f**, Robust appetitive memory implanted with 0104-GAL4 and 0273-GAL4 neuron activation contingent with odour presentation. Memory of 0104-GAL4;*UAS-dTrpA1* and 0273-GAL4;*UAS-dTrpA1* flies is significantly different from all others ($P < 0.01$, $n \geq 4$). **g**, Blocking dopamine neurons with 0104-GAL4/*UAS-shi^{ts1}* during arabinose conditioning abolishes appetitive learning ($P < 0.001$, $n \geq 8$). **h**, Blocking 0104-GAL4 dopamine neurons during sucrose conditioning significantly impairs learning ($P < 0.05$, $n \geq 6$). **i**, 0104-GAL4 neuron stimulation forms appetitive memory in satiated *Tbh* flies ($P < 0.001$, $n \geq 6$).

neither line labels dopamine neurons in the paired posterior lateral 1 (PPL1) cluster that convey negative value^{18–20} (Supplementary Fig. 3A).

We tested whether 0104-GAL4 and 0273-GAL4 PAM neurons could provide appetitive reinforcement by activating them with *UAS-dTrpA1* while presenting an odour in satiated flies. Both 0104-GAL4;*UAS-dTrpA1* and 0273-GAL4;*UAS-dTrpA1* flies exhibited robust appetitive memory that was statistically different from all control flies (Fig. 3f) and far greater than scores observed with a similar stimulation of octopamine neurons (Figs 2 and 3a).

Because 0104-GAL4 more precisely labels reinforcing PAM dopamine neurons than 0273-GAL4, we used 0104-GAL4 with *UAS-shi^{ts1}* to test whether output from PAM dopamine neurons was required for appetitive learning with sugar reinforcement. The 0104-GAL4;*UAS-shi^{ts1}* flies were tested in parallel with GAL4 driver, *UAS-shi^{ts1}* transgene and wild-type flies for comparison. Blocking 0104-GAL4 neurons abolished memory in arabinose-conditioned flies (Fig. 3g). The initial memory performance of sucrose-conditioned flies was also significantly impaired (Fig. 3h). Moreover, sucrose-conditioned 24 h memory was abolished if 0104-GAL4 neurons were only blocked during training (Supplementary Fig. 3b). Importantly, training and testing the flies at the permissive temperature did not impair performance (Supplementary Fig. 3c). Therefore PAM dopamine neurons, like octopamine neurons, are critical for conditioning with arabinose but, unlike octopamine neurons, they also contribute towards the reinforcing effects of nutritious sucrose.

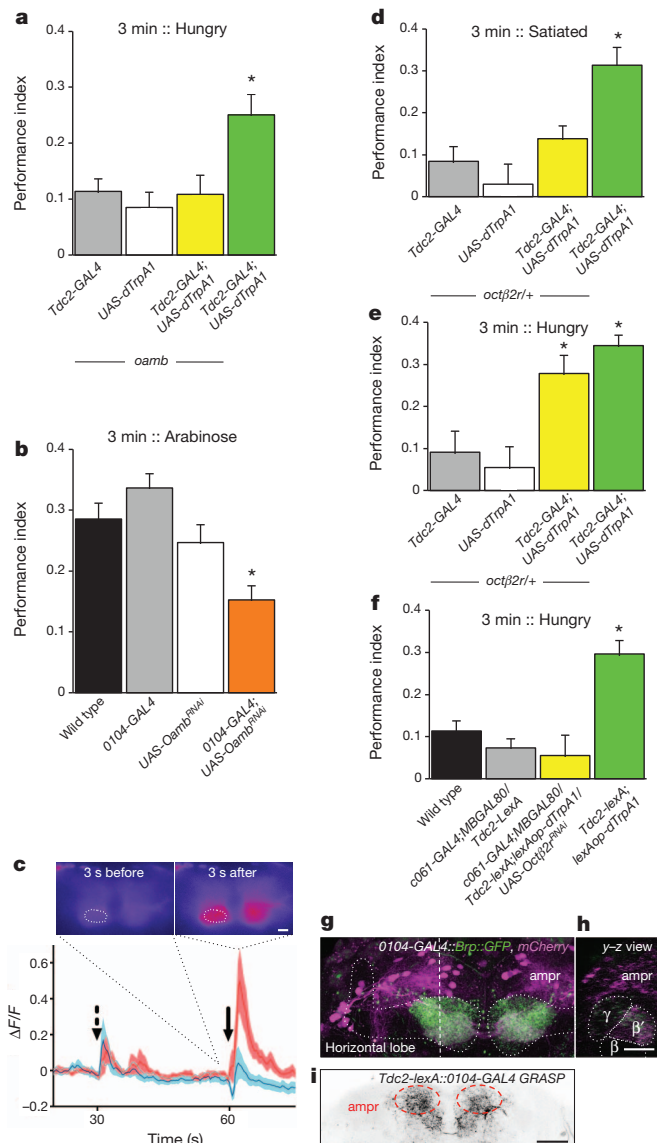
We artificially conditioned *Tbh* mutant flies that lack octopamine to investigate further whether dopamine reinforcement is downstream of octopamine. Appetitive memory formed in *Tbh* flies with 0104-GAL4;*UAS-dTrpA1* or 0273-GAL4;*UAS-dTrpA1* was statistically indistinguishable from that formed in the wild-type background (Fig. 3i and Supplementary Fig. 3d), confirming that dopamine-mediated reinforcement is downstream, and can function independently, of octopamine.

To investigate a plausible direct link between octopamine and dopamine neurons we tested whether octopamine neuron activation could form memory in octopamine receptor mutant flies. Artificial learning worked effectively in satiated *octβ1R* (also known as *oa2*) mutant flies (Supplementary Fig. 4), but was impaired in hungry *oamb* mutant flies²¹ (Fig. 4a), suggesting a key role for OAMB in reinforcement. To determine whether *Oamb* is required in PAM dopamine neurons we expressed *UAS-Oamb^{RNAi}* (Supplementary Fig. 5a) with 0104-GAL4 and conditioned flies with arabinose (Fig. 4b). The memory of 0104-GAL4;*UAS-Oamb^{RNAi}* flies was significantly different to that of both control groups. We also tested the same flies conditioned with sucrose. Consistent with previous experiments with octopamine manipulation, no effect was observed with this nutritious sugar (Supplementary Fig. 5b). Therefore, the OAMB receptor is required in PAM dopamine neurons for octopamine-dependent memory.

OAMB couples to calcium release from intracellular stores^{21,22}. We therefore expressed GCaMP3.0 (ref. 23) in PAM dopamine neurons with 0104-GAL4 and assayed intracellular Ca^{2+} responses evoked by application of exogenous octopamine. Octopamine application drove a significant increase in Ca^{2+} signal in PAM dopamine neurons that was abolished by pre-exposing the brain to the octopamine receptor antagonist mianserin^{24,25} (Fig. 4c and Supplementary Fig. 6). Therefore behavioural, anatomical and physiological data are consistent with octopamine-dependent reinforcement involving OAMB-directed modulation of PAM dopamine neurons.

Studies in *octβ2R* (ref. 24) mutant flies revealed a more nuanced picture for octopamine-mediated reinforcement. Artificial learning with octopamine neuron activation was impaired in satiated *octβ2R*/+ heterozygous flies (Fig. 4d), but was restored in *octβ2R*/+ flies by food-deprivation (Fig. 4e; note that homozygous *octβ2R* mutations are lethal). These data suggest that octopamine also integrates with systems that are responsive to hunger to provide instructive reinforcement. Such a role for octopamine is also highlighted by the observation of memory performance in all previous experiments using satiated flies (Figs 2, 3 and Supplementary Fig. 4).

Our previous work demonstrated that fly neuropeptide F (dNPF) modulates the MB–heel-innervating MB-MP1 dopamine neurons to limit retrieval of appetitive memory performance to hungry flies³. Artificial learning with octopamine worked effectively in dNPF receptor mutant flies, indicating that octopamine functions independently of dNPF (Supplementary Fig. 7). We therefore tested for a role of *Octβ2R* in MB-MP1 dopamine neurons. We used a new *Tdc2-LexA* (Supplementary Fig. 8) to simultaneously express *lexAop-dTrpA1* in



octopamine neurons and *UAS-Octβ2R^{RNAi}* in MB-MP1 neurons using *c061-GAL4;MBGAL80* (ref. 3; Fig. 4f). Hungry *Tdc2-LexA;lexAop-dTrpA1* flies formed robust appetitive memory. However, *Tdc2-LexA;lexAop-dTrpA1* flies that also carried *c061;MBGAL80;UAS-Octβ2R^{RNAi}* transgenes to knockdown *Octβ2R* expression in MB-MP1 neurons did not display memory (Fig. 4f). We tested independently the role of MB-MP1 neurons in octopamine-mediated reinforcement by simultaneously stimulating octopamine neurons while disrupting output from MB-MP1 neurons with *UAS-shi^{ts1}* (Supplementary Fig. 9). Flies in which MB-MP1 neurons were simultaneously blocked during artificial conditioning showed no significant memory. Because MB-MP1 neurons can provide aversive reinforcement if artificially engaged during odour presentation¹⁹, they probably provide a negative influence to the system²⁶. Therefore, our data indicate that octopamine-dependent appetitive reinforcement requires OCTβ2R modulation of negative dopamine signals from MB-MP1 neurons in addition to OAMB signalling in positive PAM dopamine neurons.

The *0104-GAL4* dopamine neurons have presynaptic terminals in the tip of MB β' and γ lobes and presumed dendrites in the anterior medial protocerebrum (ampr, Fig. 4g, h). We used green fluorescent protein reconstituted across synaptic partners (GRASP)^{27,28} to investigate plausible sites of synaptic contact between octopamine neurons and PAM and MB-MP1 dopamine neurons. We expressed

Figure 4 | Octopamine-dependent reinforcement functions through discrete groups of dopamine neurons. **a**, Memory cannot be implanted with *Tdc2* neuron stimulation in hungry *oamb* flies. Only *Tdc2-GAL4/UAS-dTrpA1* flies display significant learning ($P < 0.01$, $n \geq 8$). **b**, Memory formation with arabinose requires *Oamb* in 0104-GAL4 neurons. Memory of 0104-GAL4;*UAS-Oamb^{RNAi}* flies is significantly different to control groups ($P < 0.05$, $n \geq 18$). *UAS-Oamb^{RNAi}* causes ~40% decrease in *Oamb* transcript (Supplementary Fig. 5A). **c**, Applying 5 mM octopamine to the exposed fly brain drives an increase in intracellular Ca^{2+} , measured using *UAS-GCaMP3.0*, in 0104-GAL4 dopamine neurons ($\Delta F/F$ represents the evoked fluorescence change from baseline). Octopamine-evoked response (red trace) is significantly decreased in brains treated with 2.45 mM mianserin (blue trace, see also Supplementary Fig. 6). First dotted arrow; time of mianserin or vehicle application. Solid arrow; octopamine or octopamine with mianserin application. Traces averaged (each $n = 11$ flies); solid line represents mean and shaded areas s.e.m. Panels, representative pseudo-coloured images of fluorescence intensity 3 s before (left) and 3 s after (right) octopamine application. Dotted circle, analysed region of interest. Scale bar, 10 μm. **d**, Memory cannot be implanted with *Tdc2* neuron stimulation in satiated *octβ2R/+* heterozygous flies ($P > 0.05$ all groups, except *Tdc2-GAL4/UAS-dTrpA1*, $P < 0.05$, $n = 8$). **e** Memory implantation is restored in hungry *octβ2R/+* flies. *Tdc2-GAL4/UAS-dTrpA1* flies and *Tdc2-GAL4/UAS-dTrpA1; octβ2R/+* flies are significantly different to all other groups ($P < 0.05$, $n \geq 6$). **f**, Memory cannot be formed with *Tdc2* neuron stimulation in flies that express *Octβ2R^{RNAi}* in MB-MP1 neurons (all $P > 0.05$ except *Tdc2-GAL4/UAS-dTrpA1* control, $P < 0.001$, $n \geq 8$). *UAS-Octβ2R^{RNAi}* efficacy has been reported³⁰. **g**, 0104-GAL4 co-expression of mCherry (magenta) and Bruchpilot::GFP (Brp::GFP, green) reveals presynaptic label in the horizontal MB lobe tips. Brp::GFP negative processes in ampr are presumed to be PAM dopamine dendrites. **h**, Y-Z section (at the level of dashed line in **b**) reveals Brp::GFP expression only in β' and γ lobe tips. Scale bar, 20 μm. **i**, GRASP indicates contact between octopamine and 0104-GAL4 dopamine neurons in the ampr (dashed circles). Scale bar, 50 μm.

lexAop-mCD4::spGFP11 with *Tdc2-LexA* and *UAS-mCD4::spGFP11-10* with *0104-GAL4* or *c061-GAL4;MBGAL80*. Both of these combinations revealed strong GFP labelling in the ampr (Fig. 4i and Supplementary Fig. 10). In addition, MB-MP1 dopamine neuron:octopamine GRASP labelled the MB-heel region. The best candidates to bridge these two regions are the OA-VPM4 neurons, which densely innervate the MB heel and γ lobes and the ampr¹⁴. However, OA-VPM4 neurons are cleanly labelled in *0891-GAL4* (Fig. 2i) and included in the *0665-GAL4* (Fig. 2h) and *NP7088* (Fig. 2g) *GAL4* lines, all of which were insufficient for appetitive conditioning (Fig. 2j). The rest of the ampr-innervating neurons, OA-VUMa6, OA-VUMa7, OA-VUMa8 and OA-VPM3, are also included in the *NP7088 GAL4*-labelled population¹⁴. Finally, the MB-calyx-innervating OA-VPM5 neurons that are in *Tdc2-GAL4* but not *NP7088* do not have arborizations in the ampr or MB heel¹⁴, so cannot provide direct modulation of PAM or MB-MP1 dopamine neurons. Therefore, reinforcing octopamine in the fly is provided by a distributed set of neurons, some of which have arborizations in the ampr where they modulate reinforcing PAM and MB-MP1 dopamine neurons. We speculate that octopamine reinforcement may also require simultaneous regulation of other unidentified dopamine neurons, or involvement of additional parallel modes of octopamine action.

METHODS SUMMARY

Flies. All behavioural or anatomical experiments using transgenic expression were carried out on the F₁ progeny obtained from crossing *GAL4* driver flies to those harbouring reporter or effector transgenes.

Behaviour. Appetitive olfactory memory was measured as described¹¹ with modifications: for neural blockade using *UAS-shi^{ts1}*, flies were incubated at 31 °C for 30 min before and during training and testing 3 min memory. For memory implantation using *UAS-dTrpA1*, flies in filter-paper-lined tubes were presented with one odour for 2 min at 23 °C. They were then transferred into a new pre-warmed filter-paper-lined tube and presented with a second odour for 2 min at 31 °C. Flies were returned to 23 °C and tested for memory. Flies tested in the satiated condition were food-deprived 14–16 h then transferred onto fresh food for 4 h before training. Odours were 3-octanol and 4-methylcyclohexanol. Statistical analyses were performed using PRISM (GraphPad Software). Overall

analyses of variance (ANOVA) were followed by planned pairwise comparisons between the relevant groups with a Tukey honestly significant difference post-hoc test.

Immunohistochemistry. Brains were analysed for native GFP, mRFP (red fluorescent protein), mCherry, anti-TH staining, BRP::GFP and GRASP as described²⁹. Overnight preincubation in 10% NGS at 4 °C was followed by primary antiserum incubation for 3 days at 4 °C. After washing, secondary antibodies were applied for 2 days at 4 °C before brains were embedded and imaged.

In vivo calcium imaging. Fly brains expressing UAS-GCaMP3.0 (ref. 23) were imaged in saline under a SliceScope microscope (Scientifica) using a Pike CCD camera (Allied). 30 s of baseline fluorescence was recorded, followed by bath application of either saline or mianserin (Sigma-Aldrich). After another 30 s, octopamine (Sigma-Aldrich) or octopamine plus mianserin were added. Image processing, data analysis and presentation were performed with Fiji/ImageJ 1.4, Excel and Prism 6 (GraphPad Software).

Full Methods and any associated references are available in the online version of the paper.

Received 19 May; accepted 24 September 2012.

Published online 28 October 2012.

- Dayan, P. & Balleine, B. W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298 (2002).
- Wise, R. A. Dopamine, learning and motivation. *Nature Rev. Neurosci.* **5**, 483–494 (2004).
- Krashes, M. J. *et al.* A neural circuit mechanism integrating motivational state with memory expression in *Drosophila*. *Cell* **139**, 416–427 (2009).
- Liu, C. *et al.* A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* **488**, 512–516 (2012).
- Hammer, M. An identified neuron mediates the unconditioned stimulus in associative olfactory learning in honeybees. *Nature* **366**, 59–63 (1993).
- Hammer, M. & Menzel, R. Multiple sites of associative odour learning as revealed by local brain microinjections of octopamine in honeybees. *Learn. Mem.* **5**, 146–156 (1998).
- Schwaerzel, M. *et al.* Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in *Drosophila*. *J. Neurosci.* **23**, 10495–10502 (2003).
- Cole, S. H. *et al.* Two functional but noncomplementing *Drosophila* tyrosine decarboxylase genes: distinct roles for neural tyramine and octopamine in female fertility. *J. Biol. Chem.* **280**, 14948–14955 (2005).
- Monastirioti, M., Linn, C. E. J. & White, K. Characterization of *Drosophila* tyramine beta-hydroxylase gene and isolation of mutant flies lacking octopamine. *J. Neurosci.* **16**, 3900–3911 (1996).
- Kitamoto, T. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive *shibire* allele in defined neurons. *J. Neurobiol.* **47**, 81–92 (2001).
- Burke, C. J. & Waddell, S. Remembering nutrient quality of sugar in *Drosophila*. *Curr. Biol.* **21**, 746–750 (2011).
- Fujita, M. & Tanimura, T. *Drosophila* evaluates and learns the nutritional value of sugars. *Curr. Biol.* **21**, 751–755 (2011).
- Hamada, F. N. *et al.* An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* **454**, 217–220 (2008).
- Busch, S., Selcho, M., Ito, K. & Tanimoto, H. A map of octopaminergic neurons in the *Drosophila* brain. *J. Comp. Neurol.* **513**, 643–667 (2009).
- Heisenberg, M. Mushroom body memoir: from maps to models. *Nature Rev. Neurosci.* **4**, 266–275 (2003).
- Kim, Y. C., Lee, H. G. & Han, K. A. D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in *Drosophila*. *J. Neurosci.* **27**, 7640–7647 (2007).
- Gohl, D. M. *et al.* A versatile *in vivo* system for directed dissection of gene expression patterns. *Nature Methods* **8**, 231–237 (2011).
- Claridge-Chang, A. *et al.* Writing memories with light-addressable reinforcement circuitry. *Cell* **139**, 405–415 (2009).
- Aso, Y. *et al.* Specific dopaminergic neurons for the formation of labile aversive memory. *Curr. Biol.* **20**, 1445–1451 (2010).
- Mao, Z. & Davis, R. L. Eight different types of dopaminergic neurons innervate the *Drosophila* mushroom body neuropil: anatomical and physiological heterogeneity. *Front. Neural Circuits* **3**, 5 (2009).
- Han, K. A., Millar, N. S. & Davis, R. L. A novel octopamine receptor with preferential expression in *Drosophila* mushroom bodies. *J. Neurosci.* **18**, 3650–3658 (1998).
- Balfanz, S., Strunker, T., Frings, S. & Baumann, A. A family of octopamine receptors that specifically induce cyclic AMP production or Ca²⁺ release in *Drosophila melanogaster*. *J. Neurochem.* **93**, 440–451 (2005).
- Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).
- Maquieira, B., Chatwin, H. & Evans, P. D. Identification and characterization of a novel family of *Drosophila* beta-adrenergic-like octopamine G-protein coupled receptors. *J. Neurochem.* **94**, 547–560 (2005).
- Crocker, A. & Sehgal, A. Octopamine regulates sleep in *Drosophila* through protein kinase A-dependent mechanisms. *J. Neurosci.* **28**, 9377–9385 (2008).
- Waddell, S. Dopamine reveals neural circuit mechanisms of fly memory. *Trends Neurosci.* **33**, 457–464 (2010).
- Feinberg, E. H. *et al.* GFP reconstitution across synaptic partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353–363 (2008).
- Gordon, M. D. & Scott, K. Motor control in a *Drosophila* taste circuit. *Neuron* **61**, 373–384 (2009).
- Pitman, J. L. *et al.* A pair of inhibitory neurons are required to sustain labile memory in the *Drosophila* mushroom body. *Curr. Biol.* **21**, 855–861 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to M. Yoshihara, S. DasGupta, V. Budnik and S. Goodwin for reagents. We thank T. Clandinin and E. Kravitz for collegial exchange. D.O. was supported by an EMBO Long-Term Fellowship and a Sir Henry Wellcome Postdoctoral Fellowship. D.G. was supported by a Ruth L. Kirschstein NRSA Postdoctoral Fellowship (F32EY020040). M.S. was supported by a Jane Coffin Childs Postdoctoral Fellowship. S.W. is funded by a Wellcome Trust Senior Research Fellowship in the Basic Biomedical Sciences, by grants MH069883 and MH081982 from the National Institutes of Health and by funds from the Gatsby Charitable Foundation and Oxford Martin School.

Author Contributions S.W., C.J.B. and W.H. conceived this project and designed all experiments. C.J.B. and W.H. constructed fly strains, C.J.B. performed most behaviour, with some assistance from E.P. Anatomical data were produced by W.H. and C.B. Live imaging was performed by D.O. and W.H. The study was initiated by the experiments of M.J.K. G.D. constructed *lexAop-dTrpA1*. The 0104-, 0273-, 0665- and 0891-GAL4 flies were generated and initially characterized by D.G. and M.S. S.C. constructed and initially characterized *Tdc2-lexA* flies. S.W., W.H. and C.B. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.W. (scott.waddell@cnb.ox.ac.uk).

METHODS

Flies. Fly stocks were raised on standard cornmeal–agar food at 25 °C and 60% relative humidity. The wild-type strain is Canton-S. The *Tdc2-GAL4*, *NP7088*, *UAS-shi^{ts1}* (carrying insertions on the first and third chromosome) and *UAS-dTrpA1* flies are described^{8,10,13,14}. The *UAS-mCD8::GFP* and *UAS-mCD8::mCherry* strains are those in ref. 29. The *Tbh^{nM18}*, *oa2^{f02819}*, *octβ2R^{f05679}*, *oamb⁵⁸⁴* and *dumb¹* mutant strains are described^{9,16,30,31}. The MB-MP1 expressing *c061-GAL4:MBGAL80* flies are described³. The *Tdc2-lexA:VP16* transgenic line was generated by cloning the same regulatory region as described previously⁸ into the *pBS_lexA::VP16_SV40* vector³². Transgenic flies were raised by standard procedures and lines were screened for those with appropriate expression. The GRASP reporters *lexAop-mCD4::spGFP11* and *UAS-mCD4::spGFP1-10* are described²⁸. The *lexAop-dTrpA1* was constructed by subcloning a *NotI* and *XhoI* restriction-site-flanked *dTrpA1* cDNA from pOX-dTrpA1 (P. Garrity) into the pLOT transformation vector³². Transgenic flies were commercially generated (BestGene). The *UAS-Oamb^{RNAi}* (strain number 2861GD) and *UAS-Octβ2R^{RNAi}* (strain number 104524KK) were obtained from the VDRC³³. The 247-*lexA::VP16* flies are described²⁹. 0104-GAL4, 0665-GAL4, 0891-GAL4 and 0273-GAL4 flies refer to PBac[IT.GAL4]0104, PBac[IT.GAL4]0665, PBac[IT.GAL4]0891 and PBac[IT.GAL4]0273 and were generated within the framework of the InSITE project¹⁷. The *lexAop-rCD2::mRFP*, *UAS-Brp::GFP*, and *UAS-DenMark* flies are described^{34–36}. The *UAS-GCaMP3.0* flies are described²³.

Behavioural analysis. To generate flies to block or stimulate octopamine neurons we crossed *UAS-shibire^{ts1}* or *UAS-dTrpA1* female flies to *Tdc2-GAL4* males. 6–8-day-old flies were tested. To block refined subsets of octopamine neurons we crossed *UAS-shibire^{ts1}* female flies to male *NP7088/CyO*, *0665-GAL4* or *0891-GAL4/TM3* males, and only flies negative for *CyO* or *TM3* were assayed. To stimulate *Tdc2* neurons in the *Tbh* mutant background, *Tbh^{nM18}/FM7i-GFP*; *UAS-dTrpA1* females were crossed with *Tbh^{nM18}*; *Tdc2-GAL4* males and only progeny negative for *FM7i-GFP* were assayed. To stimulate *Tdc2* neurons in the *dumb¹* mutant background, *UAS-dTrpA1;dumb¹/TM3* females were crossed with *Tdc2-GAL4;dumb¹* males and only *dumb¹* homozygous flies were assayed. To stimulate 0273-GAL4 or 0104-GAL4 PAM dopamine neurons, *UAS-dTrpA1* female flies were crossed to 0273-GAL4 or 0104-GAL4/*TM6b* male flies. To block 0104-GAL4 dopamine neurons *UAS-shibire^{ts1}* females were crossed to 0104-GAL4/*TM6b* males. To stimulate 0273-GAL4 or 0104-GAL4 neurons in the *Tbh* mutant background, *Tbh^{nM18}/FM7i*; *UAS-dTrpA1* females were crossed with *Tbh^{nM18}*; 0273-GAL4/*TM6b* or *Tbh^{nM18}*; 0104-GAL4/*TM6b* males, respectively, and progeny negative for *FM7i* and *TM6b* were assayed. To stimulate *Tdc2* neurons in the *oa2*, *oamb* and *octβ2R* mutant backgrounds *UAS-dTrpA1*; *oa2^{f02819}/TM3* or *UAS-dTrpA1*; *octβ2R^{f05679}/TM6* or *UAS-dTrpA1*; *oamb⁵⁸⁴/TM6* females were crossed to *Tdc2-GAL4*; *oa2^{f02819}/TM3* or *Tdc2-GAL4*; *octβ2R^{f05679}/TM3* or *Tdc2-GAL4*; *oamb⁵⁸⁴/TM3* males, respectively. Only flies homozygous for *oa2* and *oamb* flies were assayed. The *octβ2R* insertion is homozygous lethal so only heterozygous *octβ2R/+* flies were assayed. To express *Oamb^{RNAi}* in 0104-GAL4 neurons, we crossed *UAS-Oamb^{RNAi}* [2861GD] females to 0104-GAL4/*TM6b* males and only flies lacking *TM6b* were assayed. To express *Octβ2R^{RNAi}* in MB-MP1 neurons while stimulating *Tdc2* neurons, we crossed *c061-GAL4:MBGAL80*; *Tdc2-lexA/TM3* females to *UAS-Octβ2R^{RNAi}* [104524KK]; *lexAop-dTrpA1/TM3* males. To express *shibire^{ts1}* in MB-MP1 neurons while stimulating *Tdc2* neurons, we crossed *c061-GAL4:MBGAL80*; *Tdc2-lexA/TM3* females to *UAS-shi^{ts1}*; *lexAop-dTrpA1/TM3* males. To stimulate *Tdc2* neurons in the *npfr1^{c01896}* mutant background female *UAS-dTrpA1*; *npfr1^{c01896}* flies were crossed to *Tdc2-GAL4*; *npfr1^{c01896}* males. Heterozygous control flies were generated by crossing the respective *UAS*-transgene flies with wild-type flies. For the *oa2*, *octβ2R*, *oamb*, *dumb¹* and *npfr1* experiments, *Tdc2-GAL4*; [mutant] or *UAS-dTrpA1*; [mutant] flies were flies mutant at the same locus to generate heterozygous transgene controls within the relevant homozygous mutant background. Controls for MB-MP1 neuron manipulation were generated by crossing *c061-GAL4:MBGAL80*; *Tdc2-GAL4/TM3* or *UAS-shi^{ts1}*; *lexAop-dTrpA1/TM3* or *UAS-Octβ2R^{RNAi}* [104524KK]; *lexAop-dTrpA1/TM3* females to wild-type flies.

Mixed sex populations were tested together in all behaviour experiments unless genotype required sorting single sexes. Hungry-state experiments involved food-depriving flies for 18–20 h before training in milk bottles containing a damp filter paper. To test flies in the satiated state, flies were food-deprived 14–16 h, then transferred into fresh bottles containing food to satiate 4 h before training. The olfactory appetitive paradigm was performed as described¹¹ with the following modifications: for neural blockade experiments using *UAS-shi^{ts1}*, flies were incubated at 31 °C for 30 min before and during training and testing for 3 min memory. For permissive temperature experiments flies were kept at 23 °C at all times. For memory implantation experiments using *UAS-dTrpA1*, flies were presented with one odour at the permissive temperature 23 °C for 2 min in filter-paper-lined tubes. They were then transferred into a new pre-warmed filter-paper-lined tube

and immediately presented with a second odour at the activating 31 °C for 2 min. Flies were then returned to 23 °C and tested for immediate memory. To test 3 h memory flies were trained as above and stored in plastic vials containing dampened filter paper until testing. For 24 h memory experiments, flies were trained as above and stored in food vials for 3 h followed by 21 h of food deprivation before testing. Odours were 3-octanol (9.2 μl in 8 ml mineral oil) or 4-methylcyclohexanol (18 μl in 8 ml mineral oil). The performance index was calculated as the number of flies running towards the conditioned odour minus the number of flies running towards the unconditioned odour divided by the total number of flies in the experiment. A single performance index value is the average score from flies of the identical genotype tested with each odour.

Statistical analyses were performed using PRISM (GraphPad Software). Overall analyses of variance (ANOVA) were followed by planned pairwise comparisons between the relevant groups with a Tukey honestly significant difference post-hoc test.

Imaging. To visualize native GFP, mRFP or mCherry adult female flies were collected 2–10 days after eclosion (1 day for GRASP flies) and brains were dissected in ice-cold 4% paraformaldehyde solution in PBS (1.86 mM NaH₂PO₄, 8.41 mM Na₂HPO₄, 175 mM NaCl) and fixed for an additional 60–120 min at room temperature under vacuum. Samples were washed three times 10 min with PBS containing 0.1% Triton-X100 (PBT), and twice in PBS before mounting in Vectashield (Vector Labs).

For immunohistochemistry, brains were fixed and washed as described above, followed by overnight incubation on a shaker in 10% normal goat serum (NGS) at 4 °C. For staining using the TβH antiserum³⁰, brains were dissected in ice-cold PBS, then fixed in undiluted Bouin's solution for 20 min. Samples were washed three times 10 min with PBS, then twice 10 min with PBT and incubated in 10% NGS overnight as above. The anti-TβH antiserum³⁰ was added to a final dilution of 1:300, together with anti-nc82 antibody (mouse IgG1) and anti-GFP antibody (mouse IgG2a) to a final dilution of 1:200 each and incubated for 3 days. After washing with PBT, Alexa594-coupled donkey anti-rat, Alexa488 goat anti-mouse IgG2a, and Alexa649 goat anti-mouse IgG1 antibodies were added 1:200 for 2 days, followed by washing and embedding as described above. An anti-tyrosine hydroxylase (TH) antibody raised in rabbit (AB152, Millipore) was added to a final dilution of 1:200 and kept in same conditions for another 3 days. After washing with PBT, Alexa594-coupled goat anti-rabbit antibody (A-11037, Invitrogen) was added 1:200 for one more night, followed by washing and embedding as described before.

Imaging was performed on a Zeiss LSM 5 Pascal confocal microscope and a Leica TCS SP5 X. Images were processed in AMIRA 5.2 (Mercury Systems). In some cases, debris on the brain surface and/or antennal and gustatory nerves were manually deleted from the relevant confocal sections to permit construction of a clear projection view of the z-stack.

In vivo calcium imaging. Up to 7-day-old *UAS-GCaMP3.0*; 0104-GAL4 flies were anaesthetized on ice and waxed to a custom imaging chamber. The head capsule was opened under 800 μl of sugar-free HL3-like saline³⁷, and the whole preparation transferred under a SliceScope microscope (Scientifica). Epifluorescence images were acquired using a Pike CCD camera (Allied) at a rate of three images per second at one set gain. The spontaneous baseline GCaMP3.0 response was imaged for 30 s, then either 100 μl saline or 100 μl mianserin (12.5 mM, Sigma-Aldrich, filtered) were added to the bath. After another 30 s, 100 μl octopamine (50 mM, Sigma-Aldrich) or 100 μl octopamine (50 mM) + mianserin (12 mM, filtered) were added as before, to reach a final bath concentration of 5 mM octopamine and 2.45 mM mianserin, respectively. After registration of images (StackReg plugin) a standardized region of interest (ROI) was centred within the area of the MB β' lobe tip (Fig. 4c). Image processing and analysis was performed with Fiji/ImageJ 1.4. Intensity tables were exported to Excel and the ΔF/F was calculated, with an *F* consisting of the averaged first 24 images. Traces were generated in Prism 6 (GraphPad Software). Respective peak intensities within 5 s after saline/octopamine/mianserin application were selected and compared to other groups for significant differences.

Real-time PCR. Total RNA from adult fly heads was isolated with TRIzol (Invitrogen) and cleaned with RNeasy Micro Kit (Qiagen) with DNase I treatment. RNA (200 ng) was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and random primers. The cDNA was used for quantitative real-time PCR with ABI PRISM 7000 Sequence Detection System (Applied Biosystems) with standard cycling parameters (2 min at 50 °C, 10 min at 95 °C, and 45 alternate cycles of 15 s at 95 °C and 60 s at 60 °C). The PCR mixture contained TaqMan Gene Expression Master Mix and the appropriate Gene Expression Assay (Applied Biosystems). TaqMan qPCR assays were ordered for *Oamb* (Applied Biosystems primer set annotation no. Dm02150048_m1). *Gpdh* (Applied Biosystems primer set annotation no. Dm01841185_m1) was used

as endogenous control for normalization (ΔC_t value). The decrease in expression ($\Delta\Delta C_t$ value) was calculated and transformed to the exponential scale.

30. Koon, A. C. *et al.* Autoregulatory and paracrine control of synaptic and behavioral plasticity by octopaminergic signaling. *Nature Neurosci.* **14**, 190–199 (2011).
31. Lee, H. G., Seong, C. S., Kim, Y. C., Davis, R. L. & Han, K. A. Octopamine receptor OAMB is required for ovulation in *Drosophila melanogaster*. *Dev. Biol.* **264**, 179–190 (2003).
32. Lai, S. L. & Lee, T. Genetic mosaic with dual binary transcriptional systems in *Drosophila*. *Nature Neurosci.* **9**, 703–709 (2006).
33. Dietzl, G. *et al.* A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* **448**, 151–156 (2007).
34. Lee, T. & Luo, L. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* **22**, 451–461 (1999).
35. Wagh, D. A. *et al.* Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in *Drosophila*. *Neuron* **49**, 833–844 (2006).
36. Nicolai, L. J. J. *et al.* Genetically encoded dendritic marker sheds light on neuronal connectivity in *Drosophila*. *Proc. Natl Acad. Sci. USA* **107**, 20553–20558 (2010).
37. Yoshihara, M. Simultaneous recording of calcium signals from identified neurons and feeding behavior of *Drosophila melanogaster*. *J. Vis. Exp.* **62**, e3625 (2012).

Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells

Alexej Abyzov^{1,2,3}, Jessica Mariani^{1,4*}, Dean Palejev^{1,4*}, Ying Zhang^{1,5*}, Michael Seamus Haney^{6,7*}, Livia Tomasini^{1,4*}, Anthony F. Ferrandino^{1,4}, Lior A. Rosenberg Belmaker^{1,4}, Anna Szekely^{1,5,8}, Michael Wilson^{1,2,4}, Arif Kocabas^{1,4}, Nathaniel E. Calixto^{1,4}, Elena L. Grigorenko^{1,4,9,10}, Anita Huttner^{1,11}, Katarzyna Chawarska^{1,4}, Sherman Weissman^{1,5}, Alexander Eckehart Urban^{1,6,7}, Mark Gerstein^{1,2,3,12} & Flora M. Vaccarino^{1,4,13}

Reprogramming somatic cells into induced pluripotent stem cells (iPSCs) has been suspected of causing *de novo* copy number variation^{1–4}. To explore this issue, here we perform a whole-genome and transcriptome analysis of 20 human iPSC lines derived from the primary skin fibroblasts of seven individuals using next-generation sequencing. We find that, on average, an iPSC line manifests two copy number variants (CNVs) not apparent in the fibroblasts from which the iPSC was derived. Using PCR and digital droplet PCR, we show that at least 50% of those CNVs are present as low-frequency somatic genomic variants in parental fibroblasts (that is, the fibroblasts from which each corresponding human iPSC line is derived), and are manifested in iPSC lines owing to their clonal origin. Hence, reprogramming does not necessarily lead to *de novo* CNVs in iPSCs, because most of the line-manifested CNVs reflect somatic mosaicism in the human skin. Moreover, our findings demonstrate that clonal expansion, and iPSC lines in particular, can be used as a discovery tool to reliably detect low-frequency CNVs in the tissue of origin. Overall, we estimate that approximately 30% of the fibroblast cells have somatic CNVs in their genomes, suggesting widespread somatic mosaicism in the human body. Our study paves the way to understanding the fundamental question of the extent to which cells of the human body normally acquire structural alterations in their DNA post-zygotically.

The ability to derive iPSCs from somatic cells^{5–8} has opened exciting new possibilities for the study of human development and regenerative medicine, and has enabled the effect of human genetic variation on developmental processes to be examined^{9–13}. However, all of these applications require that iPSCs, clonal cell lines each derived from one or just a few somatic cells, stably maintain the genetic background of the individual from whom they are derived. Yet, there are reports of genomic instability in stem and precursor cells, indicating that copy number variation/structural variation might arise in iPSCs, in addition to single base-pair changes^{1–4,14–16}. These variations could be caused by the de-differentiation procedures, result from extensive time in culture, or pre-exist in the somatic tissue of origin at low frequency. Emerging evidence suggests potentially widespread genomic mosaicism not only in cancer but also in somatic cell lineages, as a result of errors during DNA replication, DNA repair, mitosis and mobilization of transposable elements^{17–21}. Such a phenomenon could have far-reaching physiological consequences yet is still poorly understood and very difficult to study^{22–25}. The derivation of iPSCs offers the opportunity to analyse the genome of a single cell at high resolution and sensitivity.

Using the canonical retroviral method, we have produced 21 human iPSC (hiPSC) lines derived from skin fibroblasts collected from seven members of two families (Supplementary Fig. 1). The hiPSC lines were characterized by four sets of quality control criteria: (1) morphology; (2) expression of pluripotency factors at the protein level; (3) gene expression analyses (reverse transcription PCR (RT-PCR), microarrays, complete transcriptome by high-throughput RNA sequencing (RNA-seq)); and (4) demethylation of canonical pluripotency factor promoters (Supplementary Figs 2 and 3 and Supplementary Tables 1 and 2). This thorough evaluation (Supplementary Information) showed there was extensive similarity between our hiPSCs and human embryonic stem cells, and divergence of hiPSCs from the fibroblasts, indicating complete reprogramming. Finally, by using neuronal differentiation assays, we found that the hiPSCs exhibited comparable propensities for neural lineage differentiation (Supplementary Fig. 4).

We then generated one lane of whole-genome paired-end sequencing data on the Illumina HiSeq platform for 20 hiPSC lines, and predicted CNVs in hiPSC lines with CNVnator²⁶ (Supplementary Fig. 1b). CNVnator uses read-depth analysis and was shown to have the highest sensitivity in confirming CNVs previously discovered with arrays and fosmid sequencing²⁷. First, we discovered CNVs in fibroblast and hiPSC samples by comparison with the reference human genome, and then compared genotypes of each hiPSC line to their respective fibroblast cell population of origin to identify the variants manifested only in hiPSCs, that is, line-manifested CNVs (LM-CNVs). We were able to discover CNVs as small as 2 kilobases (kb), but the highest sensitivity was for CNVs of at least 5 kb in size (Supplementary Fig. 5). Using conservative criteria, we predicted a total of 74 LM-CNVs in all 20 lines (Supplementary Table 3), that is, just a few LM-CNVs per line. Similar numbers of LM-CNVs per line were observed for a few other hiPSC lines produced by the episomal method (Supplementary Information).

We observed positive yet non-significant correlations between the number of LM-CNVs and the passage number at which hiPSC lines were sequenced (Fig. 1a). Neither more relaxed CNV calling nor more sensitive criteria for LM-CNV identification made the correlation significant. LM-CNVs represent a small fraction of all CNVs that were discovered in hiPSC lines by comparison with the reference human genome and performing read-depth analysis at higher coverage (~20×) did not change the proportion of LM-CNVs versus the total number of CNVs (Fig. 1b). Even with sensitive criteria for LM-CNV prediction, their fraction did not exceed 17%. As a positive control and using the same approach, we compared an hiPSC line to the fibroblasts

¹Program in Neurodevelopment and Regeneration, Yale University, New Haven, Connecticut 06520, USA. ²Program in Computation Biology and Bioinformatics, Yale University, New Haven, Connecticut 06520, USA. ³Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520, USA. ⁴Child Study Center, Yale University, New Haven, Connecticut 06520, USA. ⁵Department of Genetics, Yale University, New Haven, Connecticut 06520, USA. ⁶Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, California 94305, USA. ⁷Department of Genetics, School of Medicine, Stanford University, Stanford, California 94305, USA. ⁸Department of Neurology, Yale University, New Haven, Connecticut 06520, USA. ⁹Department of Psychology, Yale University, New Haven, Connecticut 06520, USA. ¹⁰Department of Epidemiology and Public Health, Yale University, New Haven, Connecticut 06520, USA. ¹¹Department of Pathology, Yale University, New Haven, Connecticut 06520, USA. ¹²Department of Computer Science, Yale University, New Haven, Connecticut 06520, USA. ¹³Department of Neurobiology, Yale University, New Haven, Connecticut 06520, USA.

*These authors contributed equally to this work.

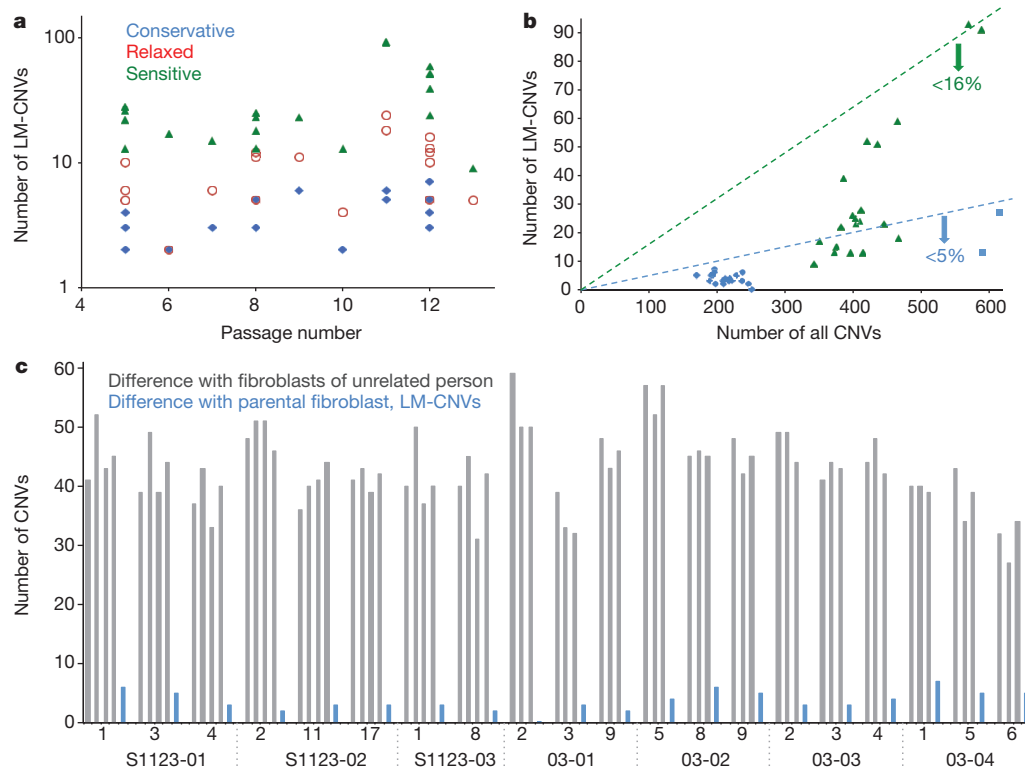


Figure 1 | Characterization of candidate LM-CNVs with respect to passage number and total CNVs. **a**, The number of LM-CNVs does not show significant changes with respect to passage, irrespective of the sensitivity of our detection criterion. Throughout this paper, conservative criteria (blue symbols) were used unless noted. **b**, The percentage of LM-CNVs of all CNVs detected in hiPSCs by comparison with the reference human genome; square symbols represent data obtained at increased (20 \times) coverage. LM-CNVs represent a small fraction the total number of CNVs in a person. **c**, Counts of LM-CNVs in

hiPSC using fibroblasts from different individuals as a baseline. The genomes of hiPSCs are different in roughly 40 CNVs (grey bars) when compared with fibroblasts from unrelated persons, that is, individuals from the other family. By contrast, genomes of hiPSCs differ by less than 10 CNVs as compared to their fibroblasts of origin (blue bars). LM-CNVs in hiPSCs as opposed to fibroblasts represent a small increment to the already existing genetic diversity in human population.

of an individual from the other family and observed roughly 40 different CNVs per comparison (that is, significantly more than the number of LM-CNVs per hiPSC line, Fig. 1c), which is consistent with inter-individual variations in a similar size range, as described previously²⁷.

Discordant paired-end reads analysis confirmed 22 LM-CNVs discovered by read-depth analysis (Supplementary information). For 39 of the most confident predictions, we performed quantitative PCR (qPCR) validation assays in early passage hiPSCs (passages 5–13), and also, when available, in late passage cells (passages 17–52) (see below). These analyses validated 33 LM-CNVs (Table 1, Supplementary Table 3 and Supplementary Figs 6–44). Validated LM-CNVs were present in 15 out of 20 (75%) hiPSC lines, with 9 (45%) hiPSC lines having more than one LM-CNV.

To obtain an independent confirmation of our approach for LM-CNV detection, we analysed the hiPSC and fibroblast samples from the mother of family S1123 and the proband of family 03 by high-resolution array-based comparative genome hybridization (aCGH). All of the 10 LM-CNVs validated by qPCR (Table 1), which were found by sequencing in the hiPSCs from these individuals, were also confirmed by aCGH (Supplementary Figs 45–54). However, no further LM-CNVs could be discovered using aCGH data, as the estimated false discovery rate (FDR) of the set of other predictions was close to 100%, based on qPCR validation of a random subset (Supplementary Tables 4 and 5). These data suggest that analysis of sequencing data alone allows the discovery of all or almost all LM-CNVs. Finally, we tested by qPCR the presence of validated LM-CNVs at later passages (passages 17–52), in five hiPSC lines. We observed a strong correlation (Pearson's coefficient 0.96) between qPCR results obtained in late versus early passages (Supplementary Fig. 6). Among 16 LM-CNVs that were tested, 87.5%

were validated in late passage (Table 1), suggesting long-term stability of the hiPSC genome.

We then analysed the origin of LM-CNVs, that is, whether they had arisen *de novo* in the hiPSCs as a sequel to reprogramming or were present at low allele frequencies in the donor fibroblast population. The first indirect, but suggestive evidence for fibroblast somatic genomic heterogeneity was the observation of the same validated LM-CNVs (chrX:64962001–65029000) in two different hiPSC lines (3 and 4) derived from the fibroblast culture of the same individual (Table 1, Fig. 2a and Supplementary Fig. 55). Further evidence for genomic heterogeneity was the realization that for many CNVs, copy number ratios were deviating from 1.5, indicative of one haplotype duplication, or from 0.5, indicative of one haplotype deletion, using both read-depth analysis and their qPCR validation (Supplementary Fig. 6 and Supplementary Information).

To test for the actual presence of somatic CNVs in the fibroblast cultures, we performed PCR amplification with diagnostic primers across CNV breakpoints in hiPSCs and the corresponding donor fibroblasts for 20 LM-CNVs with good initial estimate of their breakpoints from paired-end analyses (Fig. 2b, Table 1 and Supplementary Table 3). We observed expected bands in all cases when using hiPSC DNA, and in eight cases when using DNA from the corresponding fibroblast cultures (Supplementary Table 3; see Fig. 2b, e, g for representative examples and Supplementary Figs 7–39). For 15 LM-CNVs we also performed digital droplet PCR (ddPCR) (Fig. 2c), which allows not only the observation of low-frequency somatic CNVs but also an estimation of their allelic frequency in the somatic mosaic, with a sensitivity down to 0.1%. From the allele frequencies, cellular frequencies in the fibroblasts were calculated as explained in the Methods using the

Table 1 | Validated LM-CNVs

Person	iPSC line	LM-CNV region, type	Size (kb)	Support, no. of RPs	Somatic (PCR or ddPCR)	Frequency in fibroblasts (ddPCR)
Father S1123-01	1	chr22:38755001–38789000, dup†	34	0	–	–
	3	chr3:175005001–175064000, dup	59	3	Yes	F
		chr5:168431001–168719000, dup	288	6	UN	F
		chr20:14809001–14884000, del	75	1	Yes	F
Mother S1123-02		chrX:64962001–65029000, dup	67	3	Yes	12.6%
	4	chrX:64963001–65028000, dup	65	5	Yes	12.6%
	2	chr12:66253001–66325000, del	72	5	Yes	~0%
		chr13:111112001–111160000, del	48	3	UN	–
	11	chr4:130288001–130618000, del†	330	1	UN	F
	17	chr7:133748001–133785000, dup	37	4	Yes	14.6%
Proband S1123-03		chr11:84329001–84540000, del	211	0	–	–
		chr20:15010001–15192000, del	182	5	UN	F
	1	None	NA	NA	NA	NA
Father 03-01	8	None	NA	NA	NA	NA
	9	NA	NA	NA	NA	NA
	2	None	NA	NA	NA	NA
Mother 03-02	3	chr8:124671001–124704000, dup	33	0	–	–
		chr22:38753001–38789000, dup	36	0	–	–
	9	None	NA	NA	NA	NA
	5	chrX:90672001–90689000, del	17	2	–	–
	8	None	NA	NA	NA	NA
	9	chr1:162043001–162108000, dup	65	0	–	–
Proband 03-03		chr12:37961001–38423000, del	426	0	–	–
		chr18:70516001–70543000, del	27	0	–	–
		chrX:141153001–141191000, del	38	0	–	–
	2	chr14:76667001–76778000, del†	111	2	Yes	1.9%
		chr22:28832001–28879000, del†	47	2	UN	~0%
	3	chr5:263001–397000, del†	134	4	UN	–
Sibling 03-04	4	chr11:84581001–84688000, dup†	107	6	UN	~0%
	1	chr1:243008001–243533000, del†	525	0	–	–
		chr7:2400001–2800000, dup†	400	7	UN	–
		chr8:3558001–3685000, del†	127	3	UN	–
		chr12:37993001–38422000, del†	429	0	–	–
	5	chr1:234023001–234401000, del†	378	0	–	–
		chr8:43563001–43793000, del†	230	0	–	–
	6	chr3:143236001–143867000, dup†	631	1	Yes	0.3%
		chr8:15540001–15615000, del†	75	1	Yes	0.8%
		chr10:70514001–71136000, dup†	622	3	Yes	0.4%
		chr10:74033001–74650000, dup*†	617	3	UN	–

Further experimental support was obtained from discordantly mapped read pairs (RP) in hiPSC lines. For each CNV, no RP support was detected in the corresponding fibroblast sample. A CNV is termed 'somatic' based on either PCR or ddPCR results; UN, unknown, denotes variants whose presence could not be demonstrated in fibroblasts by either PCR or ddPCR. Zero frequency denote CNVs that are either not somatic or their frequencies are beyond the detection limit of 0.1%. Note that in a few cases a CNV was detected by PCR but its frequency could not be assessed by ddPCR. Dup, duplication; del, deletion; NA, not applicable; F, failed (events for which no successful ddPCR in both fibroblast and hiPSCs could be conducted). '–' denotes not attempted.

* Predicted dispersed duplication from paired-end analysis.

† CNV validated in late passages.

‡ CNV not validated in late passages.

ratio between the target and the control regions. The frequency of the duplication in chromosome X in fibroblast cells was estimated to be 12.6% (Fig. 2d). Cell frequencies varied from 14.6% (Fig. 2f) to less than 1% (Fig. 2h) and are summarized in Table 1 (last column). In total, the use of PCR and ddPCR allowed us to establish the presence in the parental fibroblast culture of 10 out of 20 LM-CNVs (Table 1), suggesting that fibroblast somatic genomic heterogeneity can explain at least 50% of the LM-CNVs in hiPSCs (Supplementary Table 6).

Sanger capillary sequencing of PCR bands allowed us to determine breakpoints with base-pair resolution for 18 non-redundant LM-CNVs (Supplementary Data). Analysis of sequences around breakpoints suggests non-homologous end joining as a key mechanism in the creation of LM-CNVs. Finally, we examined whether LM-CNVs affect the expression of intersected genes. Statistical analysis, using Fischer's exact test, showed that with a *P* value of 0.01 there was a direct association of gene expression with its copy number, that is, duplications increased expression whereas deletions decreased it (Supplementary Fig. 56).

In summary, we report genomic stability of hiPSC lines and the presence of extensive somatic mosaicism for copy number variation in the genome of human skin fibroblasts. This is the result of a systematic discovery and analysis of CNVs in 20 hiPSC lines relative to 7 fibroblast cultures from which the hiPSC lines were derived. As hiPSCs are clonally derived from just one or a few fibroblast cells, analysis of their genome allowed us to discover CNVs present in a subset of parental

fibroblast cells, such that very low allele frequency variants in the original populations could be unmasked. We then used PCR and ddPCR across breakpoints to genotype CNVs in the parental fibroblasts, and estimated that 50% of the CNVs manifested in hiPSCs could be traced back to the original fibroblast population (Table 1). We may be underestimating this phenomenon because very low allele frequency somatic CNVs might still escape confirmation by PCR/ddPCR in fibroblasts owing to technical limitations. Despite this, conceptually, our approach can be used for comparison of any clonal (not only iPSC) and parental cell populations with the aim of studying somatic variation.

Overall, we found that hiPSCs manifest on average two validated CNVs larger than 10 kb, which is considerably more than in two previous studies^{1,28}. The difference is probably attributable to us using sequencing (generally a more sensitive approach, see Supplementary Discussion) as opposed to single nucleotide polymorphism arrays¹. One of the aforementioned studies²⁸ also used sequencing but analysed only three hiPSC lines, therefore by extrapolating to a larger number their results could still be consistent with ours. Alternatively, bone marrow mononuclear cells may have fewer somatic variations than fibroblast cells, explaining why hiPSC lines derived from mononuclear cells in the previous study²⁸ manifest fewer LM-CNVs than do our hiPSC lines derived from fibroblasts.

It was previously proposed that CNVs might arise in hiPSCs as a consequence of DNA damage or impaired DNA repair during reprogramming. Although we acknowledge that some CNVs might arise

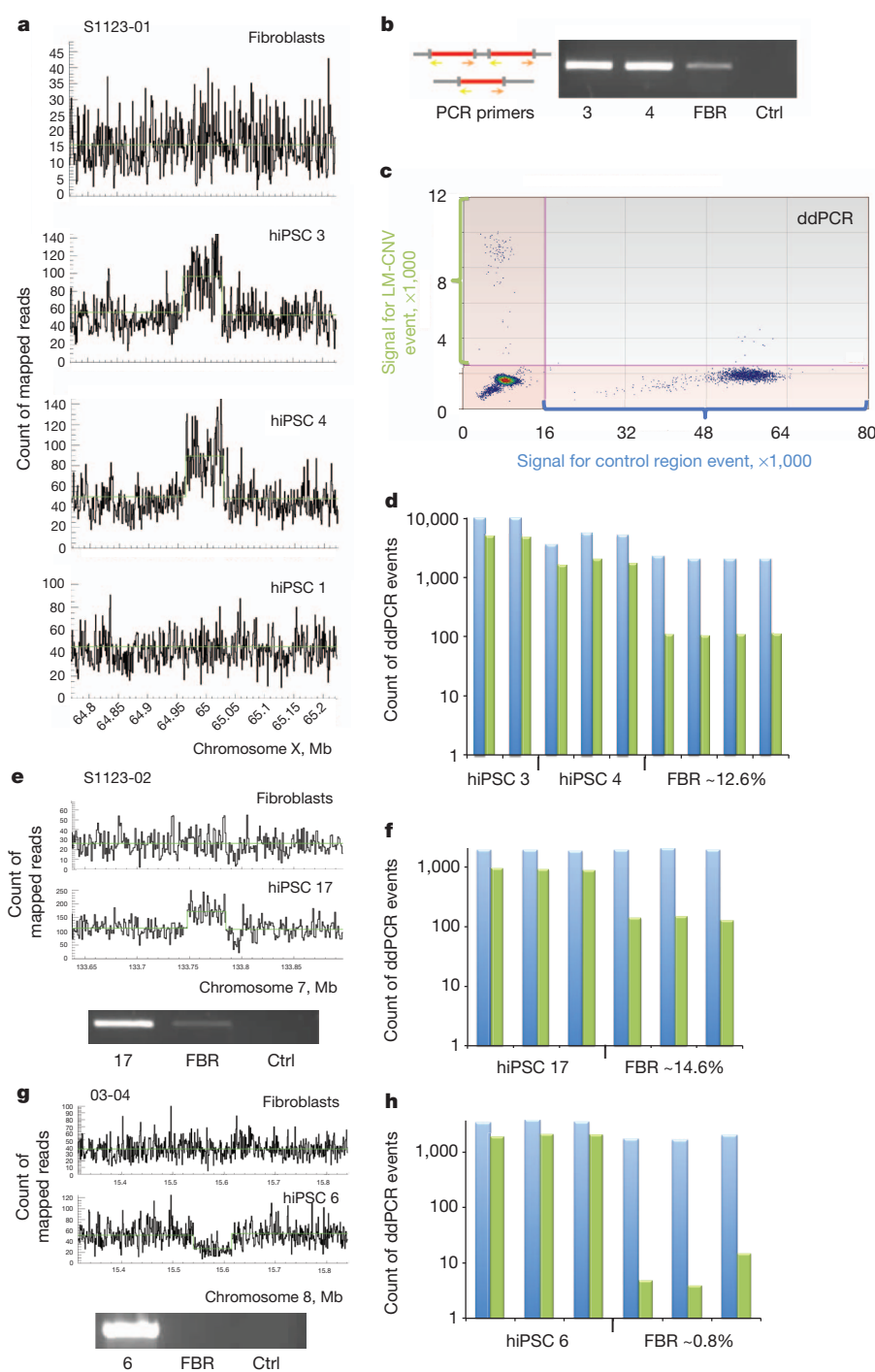


Figure 2 | Validation and estimation of cell frequency of representative somatic CNVs in fibroblasts. **a**, Two of the three hiPSC lines obtained from fibroblast sample S1123-01 had the same duplication on chromosome X not detected in parental fibroblasts. **b**, PCR amplification across CNV breakpoints showed that the duplication was present in parental fibroblasts (FBR) at a lower frequency. Ctrl, negative control. **c**, Scatter plot showing signal intensities associated with the PCR amplification across the breakpoints of the LM-CNV (y axis, green). The signal for parallel amplification of a control region is shown on the x axis (blue). Each dot represents a single PCR event. There are significantly fewer dots for PCR in CNV regions than in the control region. **d**, The frequency of cells containing the LM-CNV in fibroblasts is calculated assuming that the frequency of such cells in hiPSCs is 100%, after normalizing event numbers for LM-CNVs by the control region. Counts of ddPCR events for the LM-CNV (green bars) and the control region (blue bars) resulted in an estimated cell frequency in fibroblasts of 12.6%. **e**, Duplication on chromosome 7 that was undetectable in parental fibroblasts by read depth but detected as a faint band by PCR. **f**, This event had an estimated cell frequency in fibroblasts of 14.6% by ddPCR. **g**, Deletion on chromosome 8 that was undetectable in parental fibroblasts both by read depth and PCR. **h**, This event had an estimated cell frequency in fibroblasts of 0.8% by ddPCR.

during reprogramming in some hiPSC lines, our data suggest that reprogramming per se does not obligatorily induce *de novo* mutations, as at least half of the LM-CNVs preexisted in parental fibroblast cells (Table 1). We also found no significant difference in the number of LM-CNVs in relation to passage number. Thus, our analyses support neither the hypothesis³ that hiPSCs generally have a large rate of *de novo* mutations nor the observation that most LM-CNVs in hiPSCs disappear in late passages³. Using different parental cells and applying different protocols for cell culturing could be the factors accounting for the difference in the results.

In six hiPSC lines we determined that at least one LM-CNV originated in parental fibroblast cells. Assuming that each hiPSC colony represents a single, clonally expanded cell, we estimate that 30% (6 out

of 20) of skin fibroblast cells carry large somatic CNVs. To our knowledge, this is the first such estimate. Furthermore, with ddPCR, we estimated cell frequency as high as 15% and as low as a fraction of a per cent, suggesting wide variability in the extent of fibroblast mosaicism. Although it is possible that some CNVs could have arisen during the fibroblast cell culture²⁹, we think this is unlikely given that they were passaged less than five times (and in most cases only three times) before proceeding with hiPSC generation.

It has been known for a while²² that somatic variants can be responsible for various diseases, including cancer, and here we have provided evidence that the extent of somatic variation could have been markedly underestimated. If true, this needs to be taken into account when designing an hiPSC-based study. But more importantly, this finding

may challenge widely adopted experimental designs for genetic analyses of diseases with complex inheritance in which only the genomes of lymphoblastoid cells are being analysed. By influencing the phenotype in unexpected ways, somatically acquired CNVs might represent at least part of the explanation for the challenges in identifying the genetic contribution in some of the complex and especially in neurodevelopmental diseases, for which determining the exact loci for genetic predisposition has proven difficult³⁰.

METHODS SUMMARY

hiPSC lines were generated with Yamanaka's four retroviral vectors. One hundred nanograms of total RNA extracted from hiPSC lines was reverse-transcribed using SuperScript III Reverse Transcriptase and random hexamers. Primers used for *OCT4* (also known as *POU5F1*), *MYC* (also known as *c-Myc*) and *SOX2* specifically detect the transcripts from the endogenous genes. Neuronal differentiation was done by slight modification of an established protocol. For microarray expression analysis, total RNA was analysed by HumanHT-12 v4 BEADCHIP Illumina microarrays. Values were analysed by GenomeStudio using quantile normalization and background subtraction. RNA and DNA single-end and paired-end libraries for sequencing were prepared using standard Illumina protocols with minor modification. Whole genome sequenced reads were aligned with the reference genome with Burrows–Wheeler aligner (BWA). Primer pairs for qPCR were designed using ProbeFinder software from Roche Applied Science. Real-time qPCR was run using the Applied Biosystems StepOne Real-Time PCR System (ABI), with SYBR Green chemistry. RNA-seq reads were aligned to the reference genome with TopHat and processed with RSEQtools. Primers for PCR/ddPCR validation experiments were designed with the Primer3 software. ddPCR was performed using Quantalife machines following manufacture's instructions. For aCGH analysis, each sample was hybridized on a NimbleGen 4.2M CNV array under standard conditions as recommended by the manufacturer. Data from the arrays were analysed with Nexus Copy Number.

Full Methods and any associated references are available in the online version of the paper.

Received 16 January; accepted 28 September 2012.

Published online 18 November 2012.

- Laurent, L. C. *et al.* Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell* **8**, 106–118 (2011).
- Quinlan, A. R. *et al.* Genome sequencing of mouse induced pluripotent stem cells reveals retroelement stability and infrequent DNA rearrangement during reprogramming. *Cell Stem Cell* **9**, 366–373 (2011).
- Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471**, 58–62 (2011).
- Mayshar, Y. *et al.* Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell* **7**, 521–531 (2010).
- Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
- Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
- Wernig, M. *et al.* In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).
- Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105**, 2883–2888 (2008).
- Vaccaro, F. M. *et al.* Annual Research Review: the promise of stem cell research for neuropsychiatric disorders. *J. Child Psychol. Psychiatry* **52**, 504–516 (2011).
- Park, I. H. *et al.* Disease-specific induced pluripotent stem cells. *Cell* **134**, 877–886 (2008).
- Lee, G. *et al.* Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* **461**, 402–406 (2009).
- Hargus, G. *et al.* Differentiated Parkinson patient-derived induced pluripotent stem cells grow in the adult rodent brain and reduce motor asymmetry in Parkinsonian rats. *Proc. Natl Acad. Sci. USA* **107**, 15921–15926 (2010).
- Brennan, K. J. & Gage, F. H. Concise review: the promise of human induced pluripotent stem cell-based studies of schizophrenia. *Stem Cells* **29**, 1915–1922 (2011).
- Liang, Q., Conte, N., Skarnes, W. C. & Bradley, A. Extensive genomic copy number variation in embryonic stem cells. *Proc. Natl Acad. Sci. USA* **105**, 17453–17456 (2008).
- Wu, H. *et al.* Copy number variant analysis of human embryonic stem cells. *Stem Cells* **26**, 1484–1489 (2008).
- Elliott, A. M., Elliott, K. A. & Kammesheid, A. High resolution array-CGH characterization of human stem cells using a stem cell focused microarray. *Mol. Biotechnol.* **46**, 234–242 (2010).
- O'Huallachain, M., Karczewski, K. J., Weissman, S. M., Urban, A. E. & Snyder, M. P. Extensive genetic variation in somatic human tissues. *Proc. Natl Acad. Sci. USA* <http://dx.doi.org/10.1073/pnas.1213736109> (5 October 2012).
- De, S. Somatic mosaicism in healthy human tissues. *Trends Genet.* **27**, 217–223 (2011).
- Baillie, J. K. *et al.* Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* **479**, 534–537 (2011).
- Coufal, N. G. *et al.* L1 retrotransposition in human neural progenitor cells. *Nature* **460**, 1127–1131 (2009).
- Rehen, S. K. *et al.* Constitutional aneuploidy in the normal human brain. *J. Neurosci.* **25**, 2176–2180 (2005).
- Yousoufian, H. & Pyeritz, R. E. Mechanisms and consequences of somatic mosaicism in humans. *Nature Rev. Genet.* **3**, 748–758 (2002).
- Piotrowski, A. *et al.* Somatic mosaicism for copy number variation in differentiated human tissues. *Hum. Mutat.* **29**, 1118–1124 (2008).
- Mkrtchyan, H. *et al.* Early embryonic chromosome instability results in stable mosaic pattern in human tissues. *PLoS ONE* **5**, e9591 (2010).
- Poduri, A. *et al.* Somatic activation of AKT3 causes hemispheric developmental brain malformations. *Neuron* **74**, 41–48 (2012).
- Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011).
- Mills, R. E. *et al.* Mapping copy number variation by population-scale genome sequencing. *Nature* **470**, 59–65 (2011).
- Cheng, L. *et al.* Low incidence of DNA sequence variation in human induced pluripotent stem cells generated by nonintegrating plasmid expression. *Cell Stem Cell* **10**, 337–344 (2012).
- Arlt, M. F., Ozdemir, A. C., Birkeland, S. R., Wilson, T. E. & Glover, T. W. Hydroxyurea induces de novo copy number variants in human cells. *Proc. Natl Acad. Sci. USA* **108**, 17360–17365 (2011).
- Eichler, E. E. *et al.* Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Rev. Genet.* **11**, 446–450 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge support from the National Institutes of Health (NIH) and from the AL Williams Professorship fund and the Harris Professorship fund. We also acknowledge the Yale University Biomedical High Performance Computing Center and its support team (in particular, R. Bjornson and N. Carriero). We thank A. Klin for help with family recruitment. We thank M. V. Simonini for technical help, I.-H. Park for advice in the characterization of iPSC lines and the gift of the iPSC PGP1-1, and S. A. Duncan for the gift of the K3 iPSC line. We acknowledge the following grant support: NIMH MH089176 and MH087879, the Simons Foundation (SFARI 137055 F.V.) and the State of Connecticut, which funded the hiPSC generation and characterization; and NIH grant RR19895, which funded the instrumentation. We acknowledge the Yale Center for Clinical Investigation for clinical support in obtaining the biopsy specimens. We thank J. Overton for advice in carrying out DNA and RNA sequencing. Finally, we thank M. O'Huallachain and J. Li-Pook-Than for their advice on planning, carrying out and analysing the ddPCR experiments.

Author Contributions The authors contributed to this study at different levels, as described in the following. Study conception and design: F.M.V., A.A. and A.E.U. Family selection: E.L.G. Skin biopsy: A.S. Fibroblast culture: A.H. hiPSC generation and characterization: L.A.R.B., J.M. and L.T. Virus production: A.K. Microarrays data analysis: L.T. Neuronal differentiation: L.A.R.B., N.E.C. and J.M. Sequencing library preparation: L.A.R.B., J.M., L.T. and Y.Z. Processing and analysis of RNA-seq data: D.P. and A.A. Processing and analysis of DNaseq data: A.A. and M.W. qPCR validation: A.F.F. PCR validation: Y.Z. and A.A. aCGH hybridization and analysis: M.S.H. ddPCR experiments and analysis: M.S.H. and A.A. Human subjects: K.C. Coordination of analyses: F.M.V., S.W., A.E.U. and M.G. Display item preparation: A.A., F.M.V., L.T., D.P., J.M., N.E.C., Y.Z. and M.S.H. Manuscript writing: A.A., F.M.V. and A.E.U. The following authors contributed equally to the study: J.M., D.P., Y.Z., M.S.H. and L.T. All authors participated in discussion of results and manuscript editing.

Author Information The CNV array and sequencing data are available from Gene Expression Omnibus under accessions GSE41716 and GSE41563, and from <https://ndarportal.nih.gov/ndarportal/NDARSTD289>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.M.V. (flora.vaccaro@yale.edu), M.G. (mark.gerstein@yale.edu) or A.E.U. (aurban@stanford.edu).

METHODS

iPSC generation. A skin biopsy was obtained from the inner area of the upper arm from each member of the two families using standard techniques. Informed consent was obtained from each subject enrolled in the study according to the regulations of the Institutional Review Board and Yale Center for Clinical Investigation at Yale University. Primary cultures of fibroblasts were derived using standard procedures and infected at passage 3 with Yamanaka's four retroviral vectors, encoding for the canonical reprogramming factors (*OCT4* (also known *POU5F1*), *SOX2*, *KLF4* and *MYC*) using a multiplicity of infection (m.o.i.) of 5. After 1 month in culture, colonies with the typical human embryonic stem-cell morphology were picked, expanded on Matrigel substrate in DMEM/F12 containing 1% N2 supplement, 2% B27 supplement, 2 mM L-glutamine, 0.1 mM non-essential amino acids, 1% penicillin/streptomycin, 0.5 mg ml⁻¹ BSA fraction V (all from Invitrogen), 0.12 mM monothioglycerol (Sigma), and supplemented with 80 ng ml⁻¹ recombinant human basic fibroblast growth factor (Millipore). Colonies were characterized by immunofluorescence, RT-PCR and gene expression (see below).

RT-PCR. Total RNA was purified from hiPSC clones at passages between 5 and 13 using PicoPure RNA isolation kit (Arcturus). One hundred nanograms of total RNA extracted from hiPSC lines was reverse-transcribed using SuperScript III Reverse Transcriptase and random hexamers. Primers for embryonic stem-cell marker genes are described elsewhere³¹. Primers used for *OCT4* (also known as *POU5F1*), *MYC* and *SOX2* specifically detect the transcripts from endogenous genes. β -actin was used as a loading control.

Bisulphite sequencing. Two-hundred nanograms of genomic DNA from fibroblast cells or hiPSCs was bisulphite-converted using the MethylCode Bisulphite conversion kit (Life Technology). Bisulphite-converted DNA was amplified by PCR with the primer sets 7 for human *OCT4* (ref. 32) and sets 3 (ref. 33) for human *NANOG*. PCR was performed with the following components: 200 μ M dNTPs, 200 nM forward or reverse primer, and 2 U of PfuTurboC α hotstart DNA polymerase (Agilent Technologies), using the PCR conditions of 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 58/55 °C for 1 min and 72 °C for 1 min, followed by extension for 10 min at 72 °C. PCR products were then cloned and 7–8 colonies for each amplicon were selected for Sanger sequencing.

Neuronal differentiation. Neuronal differentiation was done by slightly modifying a protocol already used in the hiPSC field^{13,34}. Undifferentiated hiPSC colonies maintained on Matrigel were pre-incubated with the ROCK inhibitor (Y-27632), dissociated to single cells and then re-aggregated using V-bottom Aggrewell plates in serum-free medium containing recombinant noggin (200 ng ml⁻¹). After 2 days, the resulting embryoid bodies were transferred to a Petri dish, cultured in suspension for a further 2 days, and then transferred to a Matrigel substrate in serum-free medium supplemented with noggin (200 ng ml⁻¹), FGF2 (20 ng ml⁻¹) and DKK1 (200 ng ml⁻¹). After 24 h the embryoid bodies generated neuro-epithelial structures known as rosettes. A monolayer of neural progenitor cells was obtained after manual dissection, dissociation and replating of the neural rosettes on polyornithine- and laminin-coated dishes in the presence of FGF2 and EGF (both at 10 ng ml⁻¹) that allowed for the expansion (three or four passages) of the proliferating neural progenitors.

Microarrays for gene expression analysis. Total RNA isolated as above was analysed by HumanHT-12 v4 BEADCHIP Illumina microarrays. Values were analysed by GenomeStudio using quantile normalization and background subtraction. Differential scores were compared to values obtained from the federally approved H1 human embryonic stem-cell line.

Library preparations for RNA and DNA sequencing. For RNA-seq libraries, polyadenylated RNA fragments were purified by a Dynabeads mRNA Purification kit (Invitrogen), fragmented (RNA fragmentation buffer, Ambion), and reverse transcribed into first-strand complementary DNA using random hexamer and superscript II (Invitrogen), followed by second-strand cDNA synthesis using RNaseH and DNA polymerase I (Invitrogen). The cDNA was end-repaired and added a single 'A' at the 3' ends before ligating with Illumina adaptors. After running on a gel, DNA fragments from 250 to 350 bp were cut out and extracted using MinElute gel purification kit (Qiagen), and PCR-amplified using Phusion High-Fidelity master mix and Illumina primers with the condition of 98 °C for 30 s, 15 cycles of 98 °C for 10 s, 65 °C for 30 s, and 72 °C for 30 s, and concluding with 72 °C for 5 min.

To make DNA libraries, the Illumina protocol of paired-end DNA sample preparation was followed with minor modification. In short, genomic DNA was sonicated to generate fragments ranging from 200 to 800 bp, which were end-repaired, 'A' attached at the end, ligated with Illumina paired-end adaptors, size selected (450–550 bp) on 2% E-gel (Invitrogen) and extracted from the gel. The final PCR step is the same as in RNA-seq library preparation but with 18 cycles. **Conservative prediction of LM-CNVs in hiPSC.** Using Burrows–Wheeler aligner (BWA) 0.5.9-r16 (ref. 35) aligner with options '-t 4 -q 15' we have aligned genomic sequence reads to the human reference genome used by the 1000 Genomes

Project (<ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/technical/reference>), which is based on hgRC37 and included unplaced contigs. Aligned reads were paired, mapped and sorted by BWA invoked with the following options '-a 1000 -n 1 -N 1'. As a result, for each sequenced sample we obtained a file with mapped reads in BAM format. To predict CNVs, the BAM files were processed by the CNVnator method^{26,36}, which is based on read-depth analysis (see ref. 27 for review). For analysis of genomes sequenced at low coverage we used 1,000-bp bins. For analysis of two genomes sequenced at high coverage we used 400-bp bins. Then, in hiPSCs and corresponding fibroblasts, we estimated/genotyped and compared (by CNVnator) the copy number of CNVs predicted in hiPSCs. In a normal cell, the copy number should be a whole number (for example, 0, 1 or 2); however, if the population of cells used for analysis is not heterogeneous, then the copy number can be a non-negative real number (for example, 1.5). We declared copy number variation as a line-manifested deletion candidate in hiPSCs compared to fibroblasts if (1) $CN^i < 1.5$ and $CN^f > 1.5$ and $CN^f - CN^i > 0.5$; or (2) $CN^i < 0.5$ and $CN^f > 0.5$ and $CN^f - CN^i > 0.5$ for X and Y chromosomes in samples collected from males, in which CN^i and CN^f denote copy number in iPSC and fibroblast samples, respectively. Similarly, we declared copy number variation as a line-manifested duplication candidate if (3) $CN^i > 2.5$ and $CN^f < 2.5$ and $CN^i - CN^f > 0.5$; or (4) $CN^i > 1.5$ and $CN^f < 1.5$ and $CN^i - CN^f > 0.5$ for X and Y chromosomes in samples collected from males. In other words, we considered copy number variation with an estimated allele frequency in fibroblasts of at least 25% and difference in allele frequency when compared to hiPSC lines of at least 25%. We then manually inspected the read-depth signal track to select the most confident LM-CNV candidates for validation. To select confident candidates, we relied on human expertise to evaluate visually the read-depth signal in the candidate regions, presence of discordant paired-end reads supporting a prediction (see below), as well as requiring very pronounced signals in regions of segmental duplications; we also took into account whether CNVs were previously discovered CNVs^{27,37}. Two copy number variation boundaries were re-estimated. Selected confident LM-CNV candidates have been validated experimentally by qPCR, aCGH, PCR and ddPCR.

Sensitive prediction of LM-CNVs in hiPSCs. To perform a more sensitive CNV calling with CNVnator, we used option '-relax', which allowed us to find CNVs with allele frequencies down to 12.5% as opposed to 25% with the default options. Of note, the heterozygous deletion/duplications on a diploid chromosome have a 50% allele frequency. Furthermore, we relaxed the criteria on declaring a CNV as a LM-CNV. Specifically, we used the following criteria: (1) $CN^i < 1.7$ and $CN^f > 1.5$ and $CN^f - CN^i > 0.3$; and (2) $CN^i < 0.7$ and $CN^f > 0.5$ and $CN^f - CN^i > 0.3$ to call for line-manifested deletions on diploid and haploid chromosomes, respectively. Similarly, we used (3) $CN^i > 2.3$ and $CN^f < 2.5$ and $CN^i - CN^f > 0.3$; and (4) $CN^i > 1.3$ and $CN^f < 1.5$ and $CN^i - CN^f > 0.3$ to call for line-manifested duplications on diploid and haploid chromosomes, respectively. In other words, we considered CNVs with an estimated allele frequency in fibroblasts (down to 15%) and a difference in allele frequency (down to 15%) when compared to hiPSC lines.

Obtaining further support for CNVs by paired-end analysis. To obtain further support for a predicted CNV, we searched for abnormally mapped paired-ends in hiPSC lines for which CNVs were predicted and in parental fibroblasts³⁸. For a deletion, the supporting paired-ends must map with expected orientation but should have a larger span compared to the expected one from the sequencing library preparation. For a tandem duplication, the supporting paired-ends must map with an orientation different from the expected and also have a larger span (Supplementary Fig. 57). Predicted duplications may be tandem or dispersed. For dispersed duplication we searched for clusters of paired-ends with one end mapping close to predicted duplication boundaries and the other ends clustering somewhere in genome. It is well known that CNVs are enriched for repeats and homologous sequences around breakpoints, where read mapping is ambiguous³⁹. Thus, the absence of paired-end support for a predicted CNV does not invalidate the CNV. We considered a paired-end to support a deletion/duplication if it has a proper (for the type of CNV) pattern of read mapping, and its span and predicted CNV size has at least 80% mutual overlap. This condition and kilobase size of predicted CNVs guarantees that the span of supportive paired-ends is at least a few kilobases, which is much larger than the span expected from the sequencing library preparation, that is, 300–800 bp. Finally, although we did not require any particular read mapping quality, it was no less than 25 (meaning less than a 0.003 chance of incorrect mapping according to the mapper) for each supportive read. As only around 100 supportive reads were found, we do not expect any single one of them to be mapped incorrectly.

qPCR for LM-CNV call validation. Primer pairs were designed using ProbeFinder software from Roche Applied Science (<https://www.roche-applied-science.com/sis/rtprc/upl/index.jsp>). From 2 to 4 kb of DNA near the centre of the presumed CNV was scanned by ProbeFinder and the primer pair design was confirmed by UCSC In-Silico PCR (<http://genome.ucsc.edu/cgi-bin/hgPcr>) and

Primer-BLAST (<http://www.ncbi.nlm.nih.gov/tools/primer-blast>) for uniqueness and chromosomal location, only a single product and amplicon size.

The control primers to be used in reference target assays yielded a 65-bp amplicon from the *RPP30* gene (forward primer: 5'-AGATTTGGACCTGCGAGCG-3'; reverse primer: 5'-GAGCGGCTGTCTCCACAAGT-3') and a 128-bp amplicon from the *ZNF423* gene (forward primer: 5'-AGATGATCGGAGATGGTTGTG-3'; reverse primer: 5'-GATCTGCTCGTGCTCTTCAA-3'). These genes are known to be present as single copies in the haploid human genome and they have been previously used as reference in genomic qPCR assays^{40,41}. Real-time quantitative PCR was run using the Applied Biosystems StepOne Real-Time PCR System (ABI), with SYBR Green chemistry. The experimental data were processed with the StepOne Software v2.1. The comparative C_t method was used to analyse the data for the CNVs in fibroblasts and iPSCs.

All reactions for each primer set were run in triplicate and prepared from the same master mix containing 1× Power SYBR Green PCR Master Mix, 300 nM CNV forward primer, 300 nM CNV reverse primer and 10 ng genomic DNA. The thermal cycling conditions consisted of a pre-run at 95 °C for 10 min and 40 cycles with a 95 °C denaturation step for 15 s, followed by a 60 °C annealing/extension step for 60 s. The fibroblast calibrator was amplified in each run in parallel with the iPSC samples for each CNV. A no-template negative control run in duplicate was also included for each CNV assay.

RNA-seq analyses and correlation with genomic CNVs. TopHat⁴² was used to align the data against the human genome (hgRC37) and dynamically constructed exons and splice libraries. The TopHat output in BAM format was converted to SAM format using SAMtools⁴³ and then, using RSEQtools⁴⁴, to a standardized compact data format, mapped read format. For each of the GENCODE⁴⁵ genes, RSEQtools was used to compute the normalized abundance levels of transcripts measured in reads per kilobase per million mapped reads.

For each triad of hiPSCs derived from the same person, we have selected genes intersecting LM-CNVs in at least one hiPSC in the triad and having different (conservatively, more than 5 standard deviations away) from zero expression in at least one hiPSC. Then, the expression values for selected genes were compared between hiPSC in the same triad, with and without LM-CNV.

PCR to detect heterogeneity in fibroblasts. To validate LM-CNV candidates and detect heterogeneity in fibroblasts, specific primers (Supplementary Table 3) were designed to target both sides of the region adjacent to the deleted region, or the 5' and 3' ends of the duplicated region. In this way, specific products were amplified only when deletions or duplications were present. Genomic DNA from the HapMap cell line GM12878 was used as negative control. PCR was conducted with 10 ng of iPSC gDNA, 500 ng (that is, excess) of fibroblast gDNA, 500 ng of gDNA from negative control, 200 μM dNTPs, 200 nM of forward and reverse primers, 1.5 mM Mg²⁺, and 4 U of Taq polymerase (Invitrogen), using thermal cycling conditions consisting of 95 °C for 2 min, 35 cycles of 95 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s, and a final extension of 72 °C for 5 min. For one event, a second round of PCR with 30 cycles was performed to increase the signals further. For CNVs with substantial yield of PCR product in the first run, a further PCR run with 30 cycles was performed with the same conditions except reduced amounts of starting fibroblast gDNA to 10 ng (that is, equal to the amount of gDNA from hiPSC). All specifically amplified PCR bands were run on a 2% E-gel (Invitrogen), the gel was extracted by MinElute gel purification kit (Qiagen), and the extracted DNA was sequenced using both forward and reverse primers. The resulting bands were aligned to the reference genome using AGE⁴⁶ to derive the exact CNV breakpoints.

Digital PCR to estimate LM-CNV cell frequency in fibroblasts. ddPCR⁴⁷ was carried out using the Bio-Rad QX100 platform Quantalife system. Following the manufacturer's instructions, 20 μl of PCR reaction mixture consisting of ddPCR mastermix and TaqMan reagents was partitioned into 15,000–20,000 water-in-oil droplets. Each chemically homogenous droplet supports PCR amplification in a thermal cycler. TaqMan reagents enable fluorescent labelling of amplified reference and target regions. PCR products are then inserted into an automated droplet flow cytometer, where single-file, simultaneous two-colour detection of the droplets is measured. Given that the PCR mixture is randomly partitioned into 15,000–20,000 reactions vesicles, Poisson statistics can be applied to this process to yield target nucleic acid quantification of the sample.

In this instance, VIC fluorescent probes hybridizing to an amplicon targeting the *RPP30* gene served as a reference region of which two copies should be present in each cell (probes and primers provided by BioRad). LM-CNV-specific FAM probes were synthesized such that they would hybridize to amplicons targeting a given LM-CNV. Primers were designed to target LM-CNVs such that the amplicon would contain the breakpoint sequence and the FAM probe was designed to hybridize directly onto this breakpoint sequence, whenever possible (LM-CNV-specific primers and probes from IDT). In the absence of the targeted LM-CNV in

a given droplet, no PCR reaction would take place. Copy numbers of target regions were then calculated in reference to the *RPP30* event counts.

ddPCR measures allele counts of reference region and target CNVs. Let M be the measurement (that is, counts) of reference region and M_{CNV} be the measurement (that is, counts) of the target CNV allele in hiPSC. Then assuming homogeneous population of cells in hiPSC, we expect that the estimated allele frequency of a target heterozygous CNV to be ~50% for LM-CNVs on diploid chromosomes (one haplotype has no LM-CNV), and ~100% for LM-CNVs on haploid chromosomes. That is, $M_{\text{CNV}}/M \approx 0.5$ for diploid chromosomes and $(2 \times M_{\text{CNV}})/M \approx 1.0$ for haploid chromosomes (here we need to multiply by two to account for haploid chromosome, as the reference region is on diploid chromosome). Indeed we observed that measured values are very close to the mentioned expected ones, validating our assumptions that hiPSC cells are homogeneous and LM-CNVs are heterozygous.

Owing to experimental variability (for example, primer efficiency), those two ratios are slightly different from 0.5 or 1.0. Introducing as an experimental bias b accounting for the difference, then in hiPSCs, $(M_{\text{CNV}}/M) \times b = 0.5$ (for diploid chromosomes) and $(2 \times M_{\text{CNV}}/M) \times b = 1.0$ (for haploid chromosomes). Giving us $b = (0.5 \times M)/M_{\text{CNV}}$ for either diploid or haploid chromosome.

Using the same logic we can now derive an estimation of LM-CNV allele frequency in the fibroblasts. Let F be the measurement (that is, counts) of the reference allele and F_{CNV} be the measurement (that is, counts) of the target CNV allele in fibroblasts. Allele frequency can be estimated as follows: $(F_{\text{CNV}}/F) \times b$ for diploid chromosomes, and $(2 \times F_{\text{CNV}}/F) \times b$ for haploid chromosomes. b is estimated from analysis of data for hiPSC and is typically close to 1. The cell CNV frequency, that is, the fraction of cells carrying the CNV, can be estimated as $(2 \times F_{\text{CNV}}/F) \times b$ for LM-CNVs on either haploid or diploid chromosomes.

To estimate the sensitivity of the approach we performed a negative control experiment by applying primers for a LM-CNV confirmed in family S1123 to a sample from family 03, which does not have this specific LM-CNV. For 6,146 counts of reference allele in three replicates we observed only one spurious count of a LM-CNV allele. For all primers that we designed and used following the manufacturer's instruction allele ratios in hiPSC did not exceed by 16% from the expected 1:2 (one diploid chromosomes) or 1:1 (on haploid chromosomes) ratios. We thus estimate a correction factor b of less than 1.16, giving us an estimation of background noise of $(2 \times 1/6,146) \times 1.16 = 0.038\%$. Therefore, an estimation of cell frequency of 0.1% is at least 1.63 standard deviations away (assuming a Poisson nature of noise counts) from background noise.

Array CGH. Each sample was hybridized on a NimbleGen 4.2M whole-genome CNV array⁴⁸ under standard conditions as recommended by the manufacturer. Female-pooled DNA from Promega was used as the reference genome in each hybridization of the DNA samples derived from proband S1123-02. For the DNA samples derived from proband 03-03, each iPSC DNA sample was hybridized against the corresponding fibroblast DNA sample, onto the same array. After hybridization, each array was scanned on a NimbleGen MS200 Microarray scanner and the resulting images were pre-processed using NimbleScan 2.6 software. Data from the arrays were analysed further and visualized using Nexus Copy Number version 6.

Array analysis was performed in Nexus Copy Number 6 by implementing the Fast Adaptive States Segmentation Technique (FASST2) using raw probe intensity data generated by NimbleScan 2.6. This segmentation algorithm relates log₂ ratios of adjacent probes across the genome to estimate CNV events. The minimum number of probes per segment was set to 3, as this is standard for this segmentation algorithm. Thresholds for calling a copy number gain were set at a log₂ value of 0.37 and −0.5 for a copy number loss (which roughly matches the criteria of conservative calling using sequencing). Log₂ thresholds for high gains (1 or more copies) and high loss were set at 1.0 and −1.1, respectively.

Calls for proband 03-03 are candidate LM-CNVs by definition, as we hybridized hiPSC DNA against fibroblast DNA. For proband S1123-02, we selected LM-CNV candidates as calls in hiPSC that do not overlap with any call in the corresponding fibroblasts. For this person, we further filtered out calls that are likely to be noise, that is, calls smaller than 6 kb and in centromeres and telomeres.

- Chan, E. M. *et al.* Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nature Biotechnol.* **27**, 1033–1037 (2009).
- Deb-Rinker, P., Ly, D., Jezierski, A., Sikorska, M. & Walker, P. R. Sequential DNA methylation of the Nanog and Oct-4 upstream regions in human NT2 cells during neuronal differentiation. *J. Biol. Chem.* **280**, 6257–6260 (2005).
- Freberg, C. T., Dahl, J. A., Timoskainen, S. & Collas, P. Epigenetic reprogramming of OCT4 and NANOG regulatory regions by embryonal carcinoma cell extract. *Mol. Biol. Cell* **18**, 1543–1553 (2007).
- Kim, J. E. *et al.* Investigating synapse formation and function using human pluripotent stem cell-derived neurons. *Proc. Natl Acad. Sci. USA* **108**, 3005–3010 (2011).
- Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

36. Wang, L. Y., Abyzov, A., Korbelt, J. O., Snyder, M. & Gerstein, M. MSB: a mean-shift-based approach for the analysis of structural variation in the genome. *Genome Res.* **19**, 106–117 (2009).
37. Zhang, J., Feuk, L., Duggan, G. E., Khaja, R. & Scherer, S. W. Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. *Cytogenet. Genome Res.* **115**, 205–214 (2006).
38. Korbelt, J. O. *et al.* PEMer: a computational framework with simulation-based error models for inferring genomic structural variants from massive paired-end sequencing data. *Genome Biol.* **10**, R23 (2009).
39. Lam, H. Y. *et al.* Nucleotide-resolution analysis of structural variants using BreakSeq and a breakpoint library. *Nature Biotechnol.* **28**, 47–55 (2010).
40. Sanders, S. J. *et al.* Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70**, 863–885 (2011).
41. Qin, J., Jones, R. C. & Ramakrishnan, R. Studying copy number variations using a nanofluidic platform. *Nucleic Acids Res.* **36**, e116 (2008).
42. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
43. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
44. Habegger, L. *et al.* RSEQtools: a modular framework to analyze RNA-Seq data using compact, anonymized data summaries. *Bioinformatics* **27**, 281–283 (2011).
45. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
46. Abyzov, A. & Gerstein, M. AGE: defining breakpoints of genomic structural variants at single-nucleotide resolution, through optimal alignments with gap excision. *Bioinformatics* **27**, 595–603 (2011).
47. Hindson, B. J. *et al.* High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* **83**, 8604–8610 (2011).
48. Haraksingh, R. R., Abyzov, A., Gerstein, M., Urban, A. E. & Snyder, M. Genome-wide mapping of copy number variation in humans: comparative analysis of high resolution array platforms. *PLoS ONE* **6**, e27859 (2011).

Tet1 controls meiosis by regulating meiotic gene expression

Shinpei Yamaguchi^{1,2,3*}, Kwonho Hong^{1,2,3*}, Rui Liu^{4*}, Li Shen^{1,2,3}, Azusa Inoue^{1,2,3}, Dinh Diep⁴, Kun Zhang⁴ & Yi Zhang^{1,2,3,5}

Meiosis is a germ-cell-specific cell division process through which haploid gametes are produced for sexual reproduction¹. Before the initiation of meiosis, mouse primordial germ cells undergo a series of epigenetic reprogramming steps^{2,3}, including the global erasure of DNA methylation at the 5-position of cytosine (5mC) in CpG-rich DNA^{4,5}. Although several epigenetic regulators, such as Dnmt3l and the histone methyltransferases G9a and Prdm9, have been reported to be crucial for meiosis⁶, little is known about how the expression of meiotic genes is regulated and how their expression contributes to normal meiosis. Using a loss-of-function approach in mice, here we show that the 5mC-specific dioxygenase Tet1 has an important role in regulating meiosis in mouse oocytes. Tet1 deficiency significantly reduces female germ-cell numbers and fertility. Univalent chromosomes and unresolved DNA double-strand breaks are also observed in Tet1-deficient oocytes. Tet1 deficiency does not greatly affect the genome-wide demethylation that takes place in primordial germ cells, but leads to defective DNA demethylation and decreased expression of a subset of meiotic genes. Our study thus establishes a function for Tet1 in meiosis and meiotic gene activation in female germ cells.

Mouse primordial germ cells (PGCs) first appear at embryonic day (E)7.25 on the base of the allantois, and then migrate through the hindgut to the genital ridge³. During migration and at the genital ridge, PGCs undergo a series of coordinated epigenetic reprogramming, including global erasure of DNA methylation^{4,5}. Recent demonstration that proteins of the Tet family are involved in DNA demethylation prompted us to evaluate the role of Tet proteins in PGC reprogramming^{7–10}. Reverse transcriptase quantitative PCR (RT-qPCR) analysis demonstrated that *Tet1* is preferentially expressed in PGCs, whereas *Tet2* is expressed in both PGCs and somatic cells, and *Tet3* is mainly expressed in somatic cells during PGC development (E9.5–13.5) (Supplementary Fig. 1). To explore a potential role of Tet1 in PGC reprogramming and/or germ-cell development, we generated Tet1 gene-trap mice (Supplementary Fig. 2a–d). Homozygous mutant mice (*Tet1*^{Gt/Gt}) were generated by crossing heterozygous mice. Southern blot analysis, genomic sequencing and allele-specific PCR confirmed a single site insertion of a tandem-repeated gene-trap cassette into the first intron of the *Tet1* gene (Supplementary Fig. 2a–d). β -galactosidase activity analysis showed that the transgene is almost exclusively expressed in PGCs (Supplementary Fig. 2e), which is consistent with the *Tet1* expression pattern (Supplementary Fig. 1), supporting a single locus insertion. Western blot analysis demonstrated that the insertion nullified the expression of the full-length Tet1 protein (Supplementary Fig. 3a). As expected, a fusion protein between the first exon of Tet1 (amino acids 1–621) and β -geo (β -galactosidase plus neomycin resistance gene 1,303 amino acids) was detected in *Tet1*^{Gt/Gt} embryonic stem (ES) cells. Consistent with loss of Tet1, dot-blot and mass spectrometry analyses showed about a 45% reduction of 5-hydroxymethylcytosine (5hmC) levels in E9.5 *Tet1*^{Gt/Gt} embryos (Supplementary Fig. 3b–d). RT-qPCR

analysis demonstrated that the *Tet1* level in E11.5 PGCs from *Tet1*^{Gt/Gt} mice is less than 5% that of wild-type PGCs (Supplementary Fig. 3e). Consistently, Tet1 protein was also not detectable in spreads of *Tet1*^{Gt/Gt} PGCs (Supplementary Fig. 3f). Furthermore, immunostaining revealed a loss of the dotted 5hmC staining signal in germ cells of the E14.5 *Tet1*^{Gt/Gt} genital ridge (Supplementary Fig. 3g). Collectively, these data indicate that Tet1 expression is effectively abolished and the 5hmC level was significantly reduced in *Tet1*^{Gt/Gt} PGCs.

Analysis of early backcross generations (N_{1–2}) revealed they were embryonic lethal for homozygous *Tet1* mutants, whereas heterozygous mutant mice were born at normal Mendelian ratios (Supplementary Table 1). Although further backcross generations (N_{3–6}) relieved the

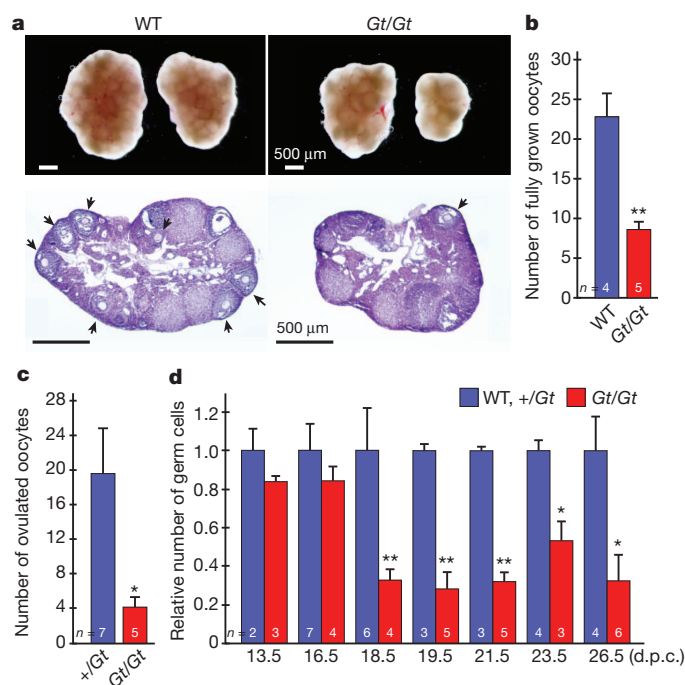


Figure 1 | The subfertility of *Tet1*^{Gt/Gt} mice is associated with oocyte loss in the late embryonic stage. **a**, Abnormal ovaries of the *Tet1*^{Gt/Gt} mice. Top, representative images of ovaries from 8-week-old wild-type (WT) and Tet1 mutant (*Gt/Gt*) mice. Both the left and right ovaries from one representative female are shown. Bottom, representative images of haematoxylin and eosin (H&E) staining of adult ovary sections. Arrows indicate fully grown oocytes. **b**, Number of fully grown oocytes in adult ovaries. *n* = 4–5. **c**, The average number of ovulated oocytes per female after hormonal stimulation. *n* = 5–7. Error bars indicate s.e.m. **d**, The relative oocyte numbers normalized to that in the control mouse. The number of oocytes in control mice was counted and set as 1. *n* = 2–7. d.p.c., days post coitum. All error bars indicate s.e.m. **P* < 0.05; ***P* < 0.01.

¹Howard Hughes Medical Institute, Harvard Medical School, WAB-149G, 200 Longwood Avenue, Boston, Massachusetts 02115, USA. ²Program in Cellular and Molecular Medicine, Boston Children's Hospital, Harvard Medical School, WAB-149G, 200 Longwood Avenue, Boston, Massachusetts 02115, USA. ³Department of Genetics, Harvard Medical School, WAB-149G, 200 Longwood Avenue, Boston, Massachusetts 02115, USA. ⁴Departments of Bioengineering, University of California at San Diego, La Jolla, California 92093-0412 USA. ⁵Harvard Stem Cell Institute, Harvard Medical School, WAB-149G, 200 Longwood Avenue, Boston, Massachusetts 02115, USA.

*These authors contributed equally to this work.

embryonic lethal phenotype, the number of viable homozygous mice is still only approximately one-third of the amount expected. Because the severity of embryonic lethality is affected by genetic background, we only used later than the N_6 generation of $Tet1^{Gt/Gt}$ mice for subsequent analyses. Similar to a recent report¹¹, reduced pup numbers were observed when either homozygous male or homozygous female animals were crossed with wild-type animals, and the pup numbers are even fewer when homozygous animals were crossed (Supplementary Fig. 4a). Because male gonad was morphologically normal and no obvious defects in male germ-cell development were observed (data not shown), our attention is focused on characterizing the female germ-cell phenotypes.

We found that the size of the $Tet1^{Gt/Gt}$ ovary is significantly smaller, with a 30% reduction in the ovary-to-body-weight ratio (Fig. 1a and Supplementary Fig. 4b). Interestingly, asymmetric ovary size caused by ovarian agenesis is frequently observed in the $Tet1^{Gt/Gt}$ animals (Fig. 1a). Both fully grown oocytes in the ovary and ovulated oocytes after hormonal stimulation are significantly reduced in the $Tet1^{Gt/Gt}$ animals (Fig. 1b, c). Ovary staining with germ-cell-specific markers (MVH, TRA98 and MSY2) followed by counting revealed a significant reduction in oocyte number from E16.5 to E18.5 (Fig. 1d and Supplementary Fig. 5a), concurrent with a significant increase in apoptotic oocytes (Supplementary Fig. 5b, c). These results suggest that increased apoptosis is probably one contributing factor for oocyte loss in $Tet1^{Gt/Gt}$ animals.

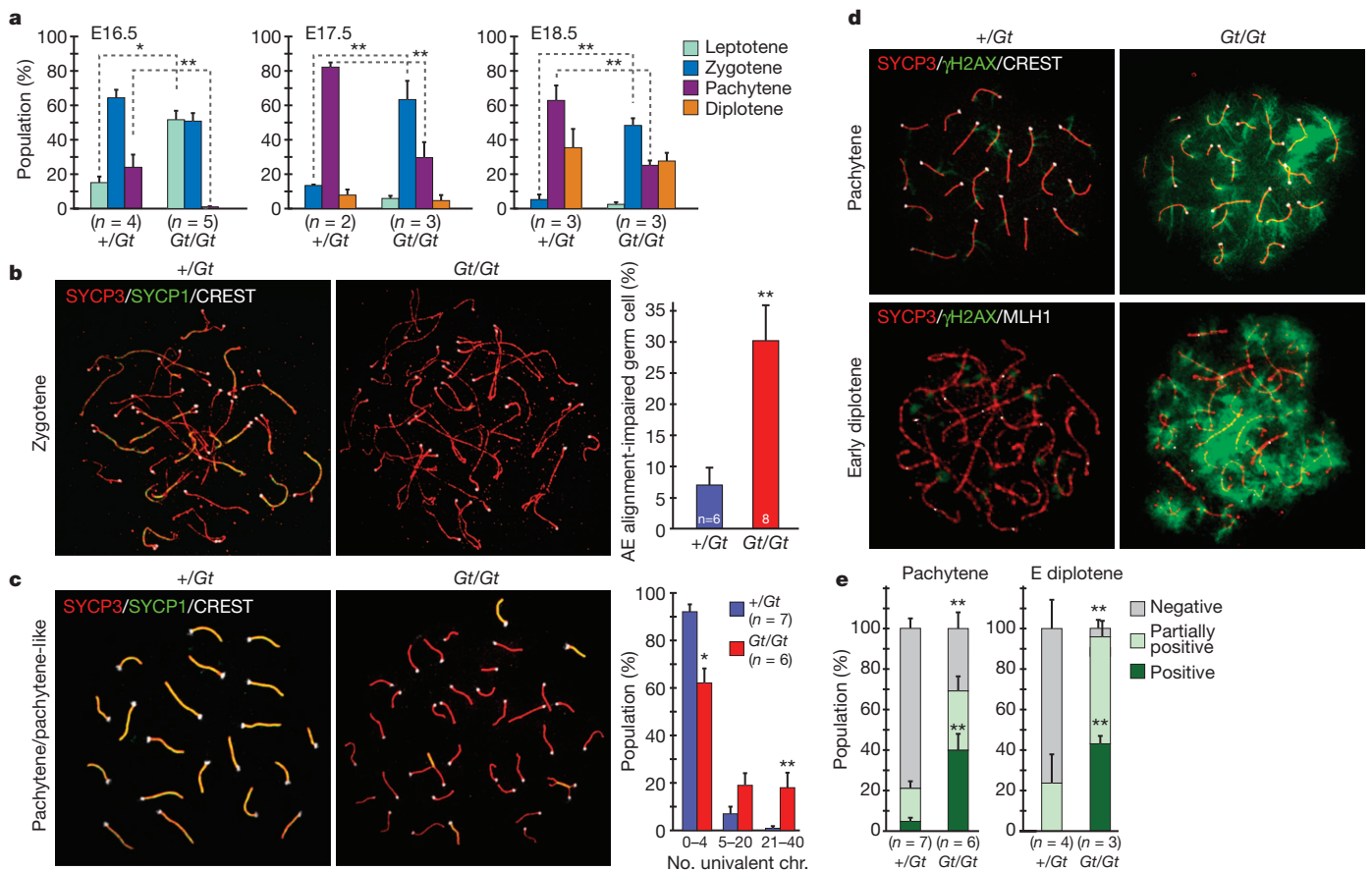


Figure 2 | Meiotic defects in $Tet1^{Gt/Gt}$ oocytes. **a**, Distribution of E16.5, E17.5 and E18.5 oocytes in the four sub-stages of meiotic prophase. $n = 2-5$. * $P < 0.05$; ** $P < 0.01$ (compared between $+/Gt$ and Gt/Gt). **b**, Left, representative images of zygotene oocytes co-stained with SYCP3, SYCP1 and CREST antibodies. Right, the percentage of oocytes with impaired axial-element alignment at the zygotene stage. Oocytes that contain less than five SYCP1 foci are counted as axial element (AE)-alignment impaired. $n = 6-8$. ** $P < 0.01$. **c**, Left, representative images of pachytene-stage oocytes co-stained with SYCP3, SYCP1 and CREST antibodies. Right, the distribution of oocyte

types categorized by the numbers of univalent chromosomes in each pachytene-stage oocyte. $n = 6-7$. * $P < 0.05$; ** $P < 0.01$. **d**, Representative images of oocytes co-stained with antibodies against γ H2AX, SYCP3 and CREST (pachytene stage), or against γ H2AX, SYCP3 and MLH1 (early diplotene stage). **e**, Distribution of oocyte types categorized by the staining pattern of γ H2AX in each pachytene- and early diplotene-stage oocyte. Representative images for each group are shown in Supplementary Fig. 9. $n = 3-7$. ** $P < 0.01$ (compared with control). All error bars indicate s.e.m.

In normal meiosis, the axial element of the synaptonemal complex starts to form at the leptotene stage, axial-element alignment and synapsis formation are initiated at the zygotene stage and completed

at the pachytene stage (Supplementary Fig. 7). However, loss of function of Tet1 resulted in an increase in the unpaired synaptonemal complex at zygotene-stage oocytes (Fig. 2b and Supplementary Fig. 7). Despite the presence of SYCP3-positive axial elements, about 30% of the zygonema contained less than four SYCP1-positive transverse filaments in *Tet1^{Gt/Gt}* oocytes (Fig. 2b), suggesting a synapsis-formation defect. Defects in synapsis formation were also observed in the pachytene-stage *Tet1^{Gt/Gt}* oocytes. In these oocytes, the vast majority of axial elements failed to align and instead remained separated from each other as univalent chromosomes (Fig. 2c and Supplementary Fig. 7). Continuous SYCP3 in short stretches of the synaptonemal complex indicates that the oocytes were at a stage corresponding to pachytene. Quantification of pachytene-stage oocytes showed that 92% of the wild-type oocytes had 0–4 univalent chromosomes, but that number dropped to 62% in the *Tet1^{Gt/Gt}* oocytes, which accompanies the increase of the oocytes with five or more univalent chromosomes (Fig. 2c). These results indicate that the Tet1 loss-of-function mutation impaired synapsis formation.

Because about 62% of pachytene-stage oocytes are successful for axial-element pairing, we asked whether they exhibited other defects. At the early phase of meiosis, DNA double-strand breaks (DSBs) are introduced in the initiation of homologous recombination, which can be detected by the presence of γ H2AX¹. As expected, immunostaining revealed the presence of γ H2AX throughout the nucleoplasm from the leptotene to zygotene stages in both wild-type and *Tet1^{Gt/Gt}* oocytes (Supplementary Fig. 8). However, although γ H2AX is gradually decreased around the pachytene stage and only a few foci remained associated with fully synapsed chromosome cores in wild-type oocytes, cloud-like nuclear staining of γ H2AX remained in the *Tet1^{Gt/Gt}* pachytene-stage and even early diplotene-stage oocytes marked by the crossover-specific marker MLH1 (ref. 13; Fig. 2d and Supplementary Fig. 8). Analysis of the γ H2AX staining pattern and quantification of the three categories (negative, partially positive and positive) (Supplementary Fig. 9) clearly demonstrated that Tet1 depletion caused accumulation of γ H2AX in pachytene- and early diplotene-stage oocytes (Fig. 2e).

The presence of γ H2AX in late-stage meiotic oocytes was also confirmed by co-staining of γ H2AX with the late-stage meiosis marker MSY2 in E18.5 ovaries (Supplementary Fig. 10). Importantly, a marked increase in γ H2AX and cleaved caspase3 double-positive cells was also observed (Supplementary Fig. 10b, c), suggesting that increased apoptotic cell death is probably caused by meiotic defects. Consistent with a DSB repair defect, the DSB repair-associated recombinase RAD51 (ref. 1) remains in pachytene and diplotene oocytes (Supplementary Fig. 11). The presence of γ H2AX and delayed removal of RAD51 in the chromosomes indicate that homologous recombination is impaired in *Tet1^{Gt/Gt}* oocytes. Staining with the crossover marker MLH1 indicated that the MLH1 foci numbers are significantly reduced in *Tet1^{Gt/Gt}* pachytene- and early diplotene-stage oocytes (Supplementary Fig. 12), further supporting a homologous recombination defect. Collectively, the above results support the notion that the Tet1 loss-of-function mutation leads to meiotic defects that include univalent chromosome formation, as well as DSB repair and homologous recombination defects.

Previous studies have shown that the establishment of proper pericentric heterochromatin (PCH) structure has an important role in meiosis¹⁴. Interestingly, immunostaining revealed specific enrichment of 5hmC at PCH of many prophase meiotic chromosomes, and this enrichment is eliminated in *Tet1^{Gt/Gt}* oocytes (Supplementary Fig. 13). As prophase meiotic PCH possesses specific histone modification patterns¹⁵, we asked whether the loss-of-function Tet1 mutation can affect PCH structure by affecting the PCH histone modification pattern. Immunostaining showed that the Tet1 loss-of-function mutation did not affect the PCH histone modification pattern (Supplementary Fig. 14), centromere clustering or localization of heterochromatin protein 1 γ (HP1 γ) (Supplementary Figs 14 and 15). Collectively, these

results suggest that the observed meiotic defect is unlikely to be due to a defect in PCH.

The stage- and cell-type-specific Tet1 expression pattern (Supplementary Fig. 1) and the dual roles of Tet1 in transcription^{16,17} suggest that the meiotic defects in *Tet1^{Gt/Gt}* germ cells might be due to aberrant transcription in PGCs. Thus, we purified PGCs from female E13.5 embryos and profiled their transcriptome by messenger RNA sequencing. We choose to focus on this time point because this is when female PGCs enter meiotic prophase after epigenetic reprogramming. We used a recently developed Smart-Seq method¹⁸ and generated more than 20 million unique reads per sample, which allowed us to identify more than 13,000 expressed transcripts in each genotype (Supplementary Table 2). Hierarchical clustering and global correlation analysis indicated that the samples were clearly separated by their genotypes, with Spearman correlation coefficients of 0.98 or 0.99 within biological replicates (Supplementary Fig. 16). Depletion of Tet1 resulted in differential expression of 1,010 genes (false discovery rate (FDR) < 0.05), among which more than 80% (899 genes) were downregulated (Fig. 3a and Supplementary Table 3). Gene Ontology (GO) analysis showed that the most significantly enriched pathways of these downregulated genes are related to the cell cycle ($P < 9 \times 10^{-11}$) and meiosis-related processes ($P < 2 \times 10^{-6}$) (Fig. 3b and Supplementary Fig. 17a). By contrast, no significant enrichment of pathways or biological processes was identified in the upregulated gene group. Importantly, genes known to be crucial for meiosis are downregulated in *Tet1^{Gt/Gt}* PGCs (Fig. 3a and Supplementary Table 3). These genes include *Stra8*, *Prdm9*, *Sycp1*, *Mael* and *Sycp3* (refs 1 and 19–22). Notably, this set of meiotic genes remained downregulated even at

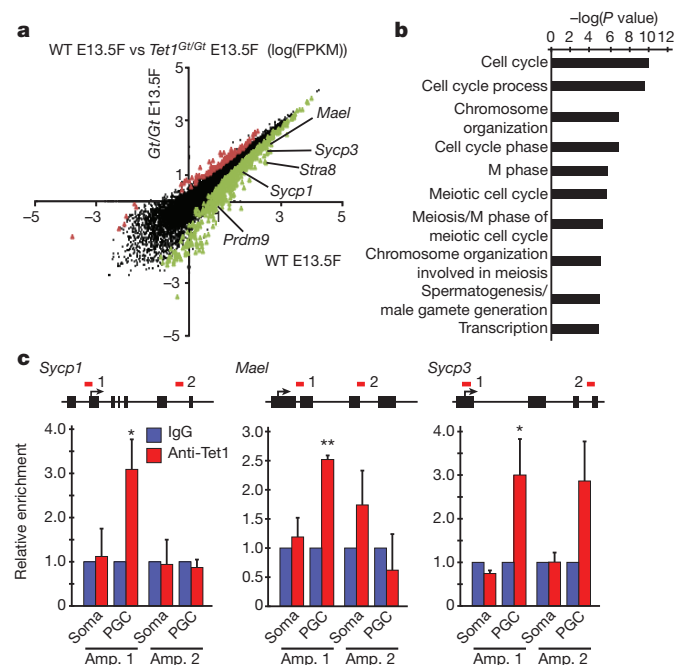


Figure 3 | Tet1 activates meiotic genes through DNA demethylation.

a, Scatter plot comparing transcriptome of wild-type and *Tet1^{Gt/Gt}* E13.5 female (E13.5F) PGCs. There are 111 and 899 genes that are respectively up- or downregulated (FDR < 0.05). Examples of downregulated meiotic genes include *Mael*, *Sycp3*, *Stra8*, *Sycp1* and *Prdm9*. FPKM, fragments per kilobase of exon per million mapped fragments. **b**, Gene Ontology analysis of downregulated genes in *Tet1^{Gt/Gt}* PGCs with a cut-off FDR < 0.05. The most enriched biological processes based on their P values are shown. **c**, ChIP-qPCR analysis of E13.5 wild-type female PGCs using an anti-Tet1 antibody demonstrates binding of Tet1 to the *Sycp1*, *Mael* and *Sycp3* promoters. Top, diagrams of *Sycp1*, *Mael* and *Sycp3* genes, with the analysed regions indicated by red lines. Bottom, relative enrichment of Tet1 compared with IgG control. Amp., amplicon. $n = 3$. Error bars indicate s.e.m. * $P < 0.05$; ** $P < 0.01$.

later developmental stages (Supplementary Fig. 17b, c), consistent with the meiotic defects observed in E16.5 *Tet1*^{Gt/Gt} PGCs. The effects of Tet1 on the expression of at least a subset of these genes are direct, as chromatin immunoprecipitation (ChIP) analysis demonstrated that Tet1 occupies the *Sycp1*, *Mael* and *Sycp3* gene promoters (Fig. 3c). Thus, Tet1 loss directly contributes to aberrant regulation of at least a subset of meiotic genes in PGCs.

To investigate how Tet1 might be involved in activation of these meiotic genes, we performed whole-genome bisulphite sequencing (WGBS) analysis using an ultra-low input, Tn5mC-seq method²³. We generated 945 million and 302 million reads for *Tet1*^{Gt/Gt} and wild-type PGCs, respectively. After removing clonal reads due to limited input cells, we obtained 14–16 million CpG sites per genotype at 1.76–2.66× genome coverage (Supplementary Table 4), which is more than 100-fold higher than a previous effort²⁴. To our knowledge, this provides the most comprehensive methylation map in PGCs so far. Consistent with previous findings²⁴, we found that PGCs are globally hypomethylated (Fig. 4a), which is verified by immunostaining (Supplementary Fig. 18). Despite no marked increase in global DNA methylation, the DNA methylation level is generally higher in mutant PGCs, particularly in exons, introns, long terminal repeats and intracisternal A particles (Fig. 4a, $P < 0.01$). With the caveat that 2× genome coverage may not allow robust identification of differentially methylated regions (DMRs) between *Tet1*^{Gt/Gt} and wild-type PGCs, we nevertheless performed detailed analysis and identified 4,337 putative

DMRs (Supplementary Table 5 and Supplementary Fig. 19) that are mostly located far away from transcriptional start sites (Supplementary Fig. 20a). These putative DMRs are associated with 5,242 genes, among which 255 exhibited altered expression (Fig. 4b and Supplementary Table 6) and are enriched for cell cycle regulation as well as for reproductive and infertility processes (FDR = 0.02) (Supplementary Fig. 20b, c). Furthermore, these DMRs are enriched for Tet1 binding in mouse ES cells¹⁷ ($P < 1 \times 10^{-100}$). To evaluate whether Tet1 binding affects DNA methylation in PGCs, we performed bisulphite sequencing on the three verified Tet1 target genes (Fig. 4c). Consistent with the involvement of Tet1 in DNA demethylation, the methylation levels of the *Sycp1*, *Mael* and *Sycp3* promoters are increased in the *Tet1*^{Gt/Gt} PGCs compared with that in the wild-type PGCs (Fig. 4c). These data indicate that Tet1-mediated demethylation of these genes is probably involved in their activation during PGC development. We note that some downregulated genes, such as *Stra8*, showed no obvious change in DNA methylation indicating that they are either regulated indirectly by Tet1 or regulated in a DNA-methylation-independent manner.

Taken together, our study provides the first evidence that Tet1 is not responsible for global demethylation in PGCs; instead, it has a specific role in meiotic gene activation at least partly through DNA demethylation. Depletion of Tet1 leads to downregulation of meiotic genes, which causes defective meiotic prophase including accumulation of non-repaired DSBs, and formation of univalent chromosomes. The meiotic defects cause loss of oocytes and consequent decrease in fertility and small litter size. Previous studies have established that the DNA methylation levels of certain meiotic genes are decreased concomitant with genomic reprogramming²⁵. Our study has extended this observation by demonstrating that Tet1 mediates locus-specific demethylation and subsequent activation of a subset of meiotic genes, revealing a specific function of Tet1 in germ-cell development.

METHODS SUMMARY

Mice and characterization of gene-trap allele. A mouse ES cell line containing *Tet1* gene-trap (Gt) allele (RRG140) was purchased from BayGenomics. To generate mice, *Tet1*^{Gt} ES cells were injected into C57BL/6J blastocysts. Heterozygous animals were backcrossed to C57BL/6J strain. Sixty micrograms of cell lysates from control and *Tet1*^{Gt/Gt} ES cells was used for western blot analysis with a Tet1 or β-galactosidase antibody. Dot-blot and mass spectrometry analyses were performed as described^{8,9}.

RNA isolation, qPCR and RNA-seq analysis using purified PGCs. To purify PGCs, the genital ridges of *Tg(Pou5f1-EGFP)* mice were dissociated with trypsin and hyaluronidase through pipetting before FACS sorting. For expression analysis, complementary DNA was synthesized by Superscript III First-Strand synthesis (Invitrogen) with random primer sets. Quantitative PCR (qPCR) reactions were performed with gene-specific primer sets (Supplementary Table 7). Relative expression levels were determined using comparative C_t values after normalizing with *Gapdh*. For RNA-seq, total RNA was purified from 500–2,000 sorted PGCs using the ZR RNA microprep kit (Zymo Research). The cDNA synthesis and amplification was performed with the SMARTer ultra low input RNA kit (Clontech). Amplified cDNAs (2–70 ng) were then fragmented by Covaris S2 sonicator and converted to sequencing libraries following the Illumina protocol for low input DNA. Bar-coded libraries were pooled and sequenced in three lanes of the Illumina HiSeq 2000 instrument.

ChIP assays and WGBS. ChIP assays were performed using the micro-ChIP method²⁶. WGBS analysis was performed using Tn5mC-seq²³. Conventional bisulphite sequencing was performed using either the EZ DNA methylation Gold kit (Zymo Research) or the EpiTect bisulphite kit (Qiagen). Primers for bisulphite sequencing are listed in Supplementary Table 7.

Full Methods and any associated references are available in the online version of the paper.

Received 14 April; accepted 25 October 2012.

Published online 14 November 2012.

- Handel, M. A. & Schimenti, J. C. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nature Rev. Genet.* **11**, 124–136 (2010).
- Hayashi, K. & Surani, M. A. Resetting the epigenome beyond pluripotency in the germline. *Cell Stem Cell* **4**, 493–498 (2009).

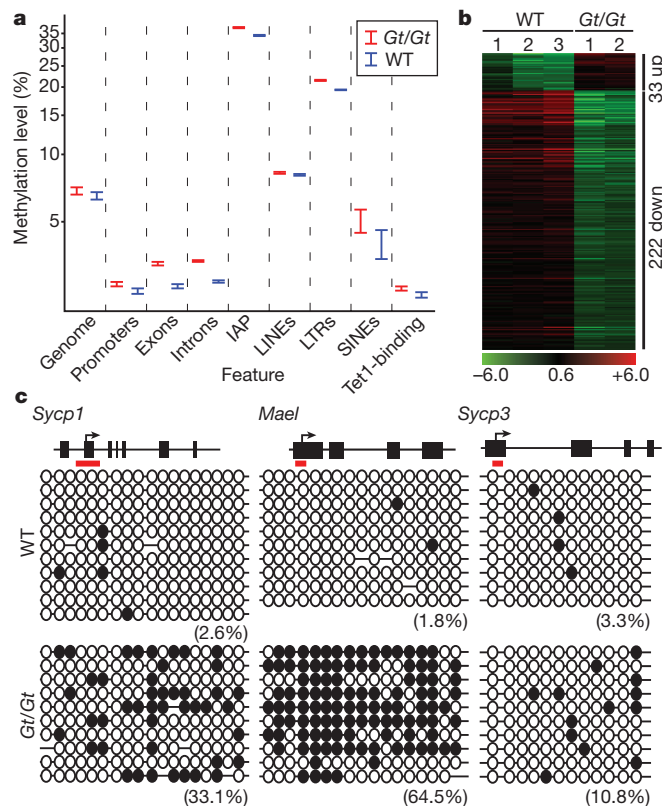


Figure 4 | WGBS analysis of the effect of *Tet1* knockout on DNA methylation in PGCs. **a**, Wild-type E13.5 female PGCs are globally hypomethylated, and depletion of Tet1 only slightly increased the global DNA methylation level. Shown are the DNA methylation levels in the entire mouse genome as well as the various genomic regions. In the last column, the DNA methylation levels at the Tet1-bound regions identified in mouse ES cells¹⁷ were compared. **b**, Heat map of the 255 differentially expressed and DMR-associated genes. **c**, Bisulphite sequencing analysis of the *Sycp1*, *Mael* and *Sycp3* gene promoters of the Tet1-binding site in wild-type and *Tet1*^{Gt/Gt} PGCs. Each CpG is represented by a circle. Open or filled circles represent unmethylated or methylated, respectively. The percentages of DNA methylation are indicated.

3. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nature Rev. Genet.* **9**, 129–140 (2008).
4. Hajkova, P. *et al.* Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).
5. Seki, Y. *et al.* Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Dev. Biol.* **278**, 440–458 (2005).
6. Kota, S. K. & Feil, R. Epigenetic transitions in germ cell development and meiosis. *Dev. Cell* **19**, 675–686 (2010).
7. He, Y. F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
8. Ito, S. *et al.* Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010).
9. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**, 1300–1303 (2011).
10. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
11. Dawlaty, M. M. *et al.* Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9**, 166–175 (2011).
12. Roeder, G. S. & Bailis, J. M. The pachytene checkpoint. *Trends Genet.* **16**, 395–403 (2000).
13. Edelmann, W. *et al.* Meiotic pachytene arrest in MLH1-deficient mice. *Cell* **85**, 1125–1134 (1996).
14. Takada, Y. *et al.* HP1 γ links histone methylation marks to meiotic synapsis in mice. *Development* **138**, 4207–4217 (2011).
15. Tachibana, M., Nozaki, M., Takeda, N. & Shinkai, Y. Functional dynamics of H3K9 methylation during meiotic prophase progression. *EMBO J.* **26**, 3346–3359 (2007).
16. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–348 (2011).
17. Wu, H. *et al.* Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature* **473**, 389–393 (2011).
18. Ramskold, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature Biotechnol.* **30**, 777–782 (2012).
19. de Vries, F. A. T. *et al.* Mouse Sycp1 functions in synaptonemal complex assembly, meiotic recombination, and XY body formation. *Genes Dev.* **19**, 1376–1389 (2005).
20. Hayashi, K., Yoshida, K. & Matsui, Y. A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* **438**, 374–378 (2005).
21. Soper, S. F. *et al.* Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev. Cell* **15**, 285–297 (2008).
22. Wang, H. & Höög, C. Structural damage to meiotic chromosomes impairs DNA recombination and checkpoint control in mammalian oocytes. *J. Cell Biol.* **173**, 485–495 (2006).
23. Adey, A. & Shendure, J. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. *Genome Res.* **22**, 1139–1143 (2012).
24. Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).
25. Maatouk, D. M. *et al.* DNA methylation is a primary mechanism for silencing postmigratory primordial germ cell genes in both germ cell and somatic cell lineages. *Development* **133**, 3411–3418 (2006).
26. Dahl, J. A. & Collas, P. A rapid micro chromatin immunoprecipitation assay (microChIP). *Nature Protocols* **3**, 1032–1045 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Wang and W. Jiang for FACS sorting, A. Adey and J. Shendure for sharing reagents and protocols for Tn5mC-seq, and H.-L. Fung for Illumina sequencing. This work was partially supported by U01DK089565 (National Institutes of Health) (to Y.Z.), R01GM097253 and CIRM BRB3-05083 (to K.Z.). Y.Z. is an Investigator of the Howard Hughes Medical Institute. D.D. is a CIRM-UCSD pre-doctoral fellow.

Author Contributions Y.Z. conceived the project; S.Y., K.H. and Y.Z. designed the experiments; S.Y., K.H., R.L., L.S. and A.I. performed the experiments; S.Y., K.H., R.L., D.D., K.Z. and Y.Z. analysed and interpreted the data; S.Y., K.H., R.L., K.Z. and Y.Z. wrote the manuscript.

Author Information RNA-seq and WGBS data have been deposited in the Gene Expression Omnibus under accession numbers GSE41908 and GSE41912, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.Z. (kzhang@bioeng.ucsd.edu) or Y.Z. (yzhang@genetics.med.harvard.edu).

METHODS

Mice and characterization of gene-trap allele. All animal studies were performed in accordance with guidelines of the Institutional Animal Care and Use Committee at the University of North Carolina at Chapel Hill. A mouse ES cell line containing the *Tet1* gene-trap (Gt) allele (RRG140) was purchased from BayGenomics. The insertion site for the gene-trap cassette was identified by Southern blotting, genomic sequencing and allele-specific PCR. Primers used for genotyping are listed in Supplementary Table 7. To generate mice, *Tet1*^{Gt} ES cells were injected into C57BL/6J blastocysts. Germ-line transmission was confirmed by PCR genotyping with allele-specific primers (Supplementary Table 7). Heterozygous animals were backcrossed to the C57BL/6J strain. Most experiments were performed using mice after backcrossing for at least six generations.

ES-cell derivation, western blot, dot-blot and mass spectrometry analysis. The *Tet1*^{Gt/Gt} ES cell line was derived as described previously²⁷ with some modifications. In brief, E3.5 blastocysts were isolated from mating of heterozygous males and females. Each blastocyst was plated on a 96-well plate with feeder layer (mitomycin C-inactivated mouse embryonic fibroblasts), and cultured with conventional ES medium (DMEM/F12 (Gibco) supplemented with 15% FBS and 1,000 U ml⁻¹ ESGRO (Millipore)). After overnight incubation, medium was changed to two inhibitors plus LIF (2iL) medium (N2B27 medium supplemented with 20% KSR (Invitrogen), 3 µM CHIR99021 (Stemgent), 15 µM PD0325901 (Stemgent) and 1,000 U ml⁻¹ ESGRO). After 5–6 days of culture, ES cells were subcloned with 0.05% Trypsin/EDTA (Gibco) and seeded onto new feeder layer. In general, ES cells are established after 5–7 passages. Genotype was determined by genomic PCR after establishment of ES lines. For western blotting analysis, feeder cells were removed and ES cells were collected after 1 day culture on the Matrigel (BD)-coated plates. Sixty micrograms of cell lysates from control and *Tet1*^{Gt/Gt} ES cells was used for western blot analysis with a Tet1 or β-galactosidase antibody. Dot-blot and mass spectrometry analyses were performed as previously described^{8,9}.

RNA isolation and qRT-PCR from purified PGCs and somatic cells. The transgenic mouse line containing the GOF18ΔPE-EGFP (*Tg(Pou5f1-EGFP)*) gene was purchased from Jackson Laboratory. Genital ridge cells were dissociated with 0.05% trypsin/EDTA treatment with pipetting. After washing with phosphate buffer medium (PB1)/BSA solution, tissues were incubated in hyaluronidase and then passed through a cell strainer. Germ cells and somatic cells were purified based on the expression of green fluorescent protein (GFP) using FACS Aria II flow cytometry (BD Bioscience). Total RNA was isolated using the RNeasy mini kit (Qiagen), and cDNA was generated with random primer sets and Superscript III first-strand synthesis system (Invitrogen). Real-time qPCR reactions were performed on an ABI ViiA7 sequence detection system (Applied Biosystems) using SYBR Green (Applied Biosystems). Relative gene expression levels were analysed using comparative C_t methods, in which C_t is the cycle threshold number, and normalized to *Gapdh*. qRT-PCR primers are listed in Supplementary Table 7.

Immunocytological analysis of spreads. To prepare the spreads, dissected E11.5 genital ridges were dissociated by 0.05% trypsin/EDTA. After centrifugation, cells were suspended in PBS and dropped onto slide glasses dipped into fixative (1% paraformaldehyde, 0.15% Triton X-100 and 3 mM dithiothreitol, pH 9.2). The slide glasses were kept overnight in a humidified box at 4 °C. The slides were washed in water containing 0.4% PhotoFlow (Kodak), and completely dried at room temperature. Surface spreads of ovary cell suspension were prepared as described²⁸ with some modifications. In brief, the ovary was minced with two forceps and dissociated by pipetting in PBS. An equal volume of hypotonic buffer (30 mM Tris-HCl, pH 8.3, 5 mM EDTA, 1.7% sucrose and 0.5% trisodium citrate dihydrate) was added to the cell suspension. After 7 min of incubation, cells were centrifuged for 5 min and resuspended in 100 mM sucrose. The cell suspension was spread onto glass slides dipped into fixative with the same method as PGC spreads.

To stain spreads, dried slide glasses were washed with 0.1% Triton X-100/PBS (PBST) for 10 min, and incubated with blocking buffer (3% BSA, 2% donkey serum/PBST) for 1 h at room temperature. For the staining with anti-5mC and -5hmC antibodies, spreads were treated with HCl solution for 20 min at room temperature, followed by washes in PBST. Spreads were then incubated with blocking buffer and primary antibodies. Spreads were incubated with primary antibody at 4 °C overnight, followed by washes with PBST, and incubation with appropriate secondary antibodies for 1 h at room temperature.

H&E and immunostaining analysis. Paraffin-embedded ovary and testis samples collected from 8-week-old mice were cut to 7-µm thickness and stained with H&E. For counting, serial sections were stained with H&E, and the oocyte numbers were manually counted at 100-µm intervals throughout the entire ovary. For immunofluorescence staining, tissues fixed with 4% paraformaldehyde/PBS overnight at 4 °C were embedded in OCT compound (Sakura) and cut to 10-µm thickness. After washing with PBS, sections were permeabilized with 0.4% Triton X-100/PBS for 20 min at room temperature, followed by 1-h incubation with blocking buffer (3% BSA, 2% donkey serum/PBST) at room temperature. Then, sections were

incubated at 4 °C with primary antibody overnight, followed by PBST washes and incubation with appropriate secondary antibodies at room temperature for 1 h. When sections were stained with anti-5hmC, sections were treated with HCl solution (4 N HCl, 0.1% Triton-X 100 in distilled water) for 20 min at room temperature, followed by washing, and incubation with blocking buffer and primary antibodies. To count the oocyte number, one in five consecutive sections was taken at regular interval throughout the entire gonad, stained with germ-cell marker and analysed by ImageJ software.

Antibodies. The antibodies used in this study include anti-SYCP3 (rabbit, Abcam ab15093), anti-SYCP3 (mouse, Abcam ab97672), anti-SYCP3 (goat, Santa Cruz sc-20845), anti-γH2AX (mouse, Millipore 05-636), anti-γH2AX (rabbit, Cell Signaling 2577), anti-CREST (human, Antibodies 15-235), anti-RAD51 (rabbit, Calbiochem PC130), anti-RAD51 (mouse, Abcam sb88572), anti-MLH1 (mouse, BD Pharmingen 550838), anti-SYCP1 (rabbit, Abcam ab15090), anti-5hmC (Active Motif 39769), anti-Tet1 (for western blotting, see ref. 8; for immunostaining, Millipore 09-872), anti-cleaved caspase3 (rabbit, Cell Signaling 9661), anti-SSEA1 (mouse, Millipore MAB4301), anti-MVH (rabbit, Abcam ab13840), anti-TRA98 (rat, BioAcademia 73-003), anti-MSY2 (goat, Santa Cruz sc-21314), anti-histone H3K9me2 (mouse, Millipore 05-1249), anti-histone H3K9me3 (rabbit, Abcam ab8898), anti-histone H4K20me3 (mouse, Abcam ab78517) and anti-HP1γ (goat, Abnova PAB6885).

Staging of meiotic germ-cell prophase. Staging of oocytes was performed based on the appearance of axial elements (SYCP3 staining) and centromere numbers (CREST staining)²⁹. In brief, in wild-type the leptotene stage is defined by the appearance of dot-like SYCP3 signals and clustering of centromeres into several loci in the nucleus; the zygotene stage is defined as the start of synapsis, and fibre-like SYCP3 staining and pairing of 30–80% of centromeres are observed; in the pachytene stage all axial elements are paired and 20 foci of CREST are visible at the end of the SYCP3 stretch; the diplotene stage is defined by more than 20 CREST signals, coincident with desynapsis, and fragmentation of SYCP3 is also observed. Because female germ cells enter meiosis in a synchronous manner, leptotene- and zygotene-stage germ cells can be obtained from E16.5 ovary, whereas pachytene and diplotene germ cells are obtained from the E18.5 ovary in the wild type. In *Tet1*^{Gt/Gt} ovaries, leptotene-stage germ cells were obtained from the E16.5 ovary, whereas zygotene, pachytene and diplotene germ cells were obtained from E18.5 ovaries. Pachytene-like stage oocytes in the *Tet1*^{Gt/Gt} ovary were defined as those that contain condensed and short axial element without any nick, no clear clustering of centromeres, and more than 20 SYCP3 stretches.

RNA-seq and GO analysis. Total RNA was purified from 500–2,000 sorted PGCs using the ZR RNA microprep kit (Zymo Research). The cDNA synthesis and amplification was performed with the SMARTer ultra low input RNA kit (Clontech). The amplified cDNA (2–70 ng) was then fragmented by Covaris S2 sonicator (Covaris) and converted to sequencing libraries following the Illumina construction protocol for low input DNA (Illumina). Bar-coded libraries were pooled and sequenced in three lanes of the Illumina HiSeq 2000 instrument.

mRNA-seq reads generated from each sample were aligned to the mouse genome (mm9, NCBI build 37) with Bowtie/Tophat v1.3.1 (<http://tophat.cbcb.umd.edu>), which allows mapping across splice sites by reads segmentation. All programs were used with default setting unless otherwise specified. Mapped reads (83–84% of total reads) were subsequently assembled into transcripts guided by reference annotation (mm9, USCS gene annotation) with Cufflinks v1.0.3 (<http://cufflinks.cbcb.umd.edu>). The expression level of each transcript was quantified with normalized FPKM. Two or three biological replicates were used in each genotype to identify transcripts that showed significant differences at a cut-off FDR < 0.05 between wild type and *Tet1*^{Gt/Gt} by Cuffdiff v1.0.3. Functional annotation of significantly different transcripts and enrichment analysis was performed with DAVID (<http://david.abcc.ncifcrf.gov>).

ChIP assays and bisulphite sequencing. ChIP assays were performed using the micro-ChIP method²⁶. In brief, GFP-positive and -negative cells from E13.5 genital ridges were cross-linked with a final concentration of 1% formaldehyde soon after sorting. After incubation at room temperature for 8 min, the reaction was stopped by the addition of 125 mM glycine. Sonicated chromatin was subjected to incubation with anti-Tet1-coated Dynabeads Protein A complex at 4 °C for 2 h. After extensive washing, immunoprecipitated DNA was eluted from the beads, and analysed on an ABI ViiA7 sequence detection system (Applied Biosystems) using SYBR Green (Invitrogen). Primers for the ChIP assay are listed in Supplementary Table 7.

Bisulphite sequencing was performed using a protocol associated with either the EZ DNA methylation Gold kit (Zymo Research) or the EpiTect bisulphite kit (Qiagen). In brief, about 20 ng of sodium bisulphite-treated DNA samples was subjected to PCR amplification using the first set of primers; PCR products were used as templates for a subsequent PCR reaction using nested primers. The PCR products of the second reaction were then subcloned using the Invitrogen TA

cloning kit following the manufacturer's instruction. PCRs and subcloning were performed in duplicate for each sample. The clones were sequenced using the M13 forward primer. Primers for bisulphite sequencing are listed in Supplementary Table 7.

WGBS and data analysis. WGBS analysis was performed using Tn5mC-seq²³. In brief, genomic DNA was isolated using the Wizard kit (Promega) when >5,000 cells were available, whereas samples with <5,000 cells were directly lysed in a hypotonic lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl and 3 mM MgCl₂) with 1% NP40 and 0.2 arbitrary units (AU) ml⁻¹ protease (Qiagen). Resultant DNA were fragmented and attached with a 5'-methylated adaptor by Tn5-transposon-based tagmentation. A 3'-methylated adaptor was added during gap-filling and ligation. Tagged DNA fragments were bisulphite-converted using the Imprint kit (Sigma) or Lightning kit (Zymo), followed by PCR amplification with Illumina adaptors. Bar-coded libraries were pooled and sequenced with Illumina HiSeq 2000 instrument.

Bisulphite-converted sequencing data were processed using the BisReadMapper analysis pipeline³⁰. In brief, reads were aligned to the mouse reference genome (mm9) using SOAP2. Libraries processed with the Sigma Imprint kit contain a mixture of fully converted and poorly converted reads. A post-mapping filtering was performed to remove unconverted reads based on the criterion that cytosine in the non-CpG context is present at <1.2% in the raw sequencing reads, which ensured a bisulphite conversion rate of >98% in the filtered reads, as estimated by the percentage of unmethylated cytosines on the mitochondrial chromosome. Libraries processed with the Zymo lightning kit were fully converted, such that the post-mapping filtering was not necessary. Data from the two batches of libraries were pooled as global analysis on the two batches yielded consistent results. The global methylation distribution plot was generated using the average methylation levels of CpGs in 10-kb non-overlapping bins along the genome. DMRs between genotypes were identified by comparing the fraction of methylated cytosines in a sliding window that contains at least six CpG sites based on

chi-square test. Repetitive regions were masked before DMR calling. Adjacent candidate DMRs (<250 base pairs (bp) apart) were joined, and the flanking non-informative CpG sites were further trimmed. We required that there are at least 20 methylated or unmethylated cytosine observations for each sample in a DMR, and the mean methylation difference is at least 0.2. The FDR was empirically estimated to be ~10%, by applying the same procedure on randomly permuted data sets in which the methylation level of each CpG site was randomly assigned to a different CpG site in the genome while maintaining the global distribution of CpG methylation in each sample. Identified DMRs were correlated with Tet1-binding sites defined in ESCs¹⁷ by requiring >1 bp overlapping using BEDtools (<http://code.google.com/p/bedtools/>). The FDR was then tested by 100 permutations of randomly placed DMRs along the genomes. DMRs were assigned to potential target genes and functional characterized by GREAT (<http://great.stanford.edu/public/html/splash.php>). For visualization, locally smoothed methylation levels were calculated by the Bsseq package with parameter of $ns = 40$, $h = 2,000$ (in which ns is the minimum methylation loci in a smoothing window, and h is the minimum smoothing window size in bases) to obtain high confidence methylation levels for each genotypes (<http://www.bioconductor.org/packages/devel/bioc/html/bsseq.html>). We did not use Bsseq for DMR calling owing to a high false positive rate on the low-coverage data we have.

27. Markoulaki, S., Meissner, A. & Jaenisch, R. Somatic cell nuclear transfer and derivation of embryonic stem cells in the mouse. *Methods* **45**, 101–114 (2008).
28. Soper, S. F. C. *et al.* Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis. *Dev. Cell* **15**, 285–297 (2008).
29. Kouznetsova, A., Novak, I., Jessberger, R. & Hoog, C. SYCP2 and SYCP3 are required for cohesin core integrity at diplotene but not for centromere cohesion at the first meiotic division. *J. Cell Sci.* **118**, 2271–2278 (2005).
30. Diep, D. *et al.* Library-free methylation sequencing with bisulfite padlock probes. *Nature Methods* **9**, 270–272 (2012).

Structure of the Mediator head module

Laurent Larivière^{1*}, Clemens Plaschka^{1*}, Martin Seizl¹, Larissa Wenzek¹, Fabian Kurth^{1†} & Patrick Cramer¹

Gene transcription by RNA polymerase (Pol) II requires the coactivator complex Mediator. Mediator connects transcriptional regulators and Pol II, and is linked to human disease^{1–4}. Mediator from the yeast *Saccharomyces cerevisiae* has a molecular mass of 1.4 megadaltons and comprises 25 subunits that form the head, middle, tail and kinase modules^{5–7}. The head module constitutes one-half of the essential Mediator core⁸, and comprises the conserved⁹ subunits Med6, Med8, Med11, Med17, Med18, Med20 and Med22. Recent X-ray analysis of the *S. cerevisiae* head module at 4.3 Å resolution led to a partial architectural model with three submodules called neck, fixed jaw and moveable jaw¹⁰. Here we determine *de novo* the crystal structure of the head module from the fission yeast *Schizosaccharomyces pombe* at 3.4 Å resolution. Structure solution was enabled by new structures of Med6 and the fixed jaw, and previous structures of the moveable jaw¹¹ and part of the neck¹², and required deletion of Med20. The *S. pombe* head module resembles the head of a crocodile with eight distinct elements, of which at least four are mobile. The fixed jaw comprises tooth and nose domains, whereas the neck submodule contains a helical spine and one limb, with shoulder, arm and finger elements. The arm and the essential shoulder contact other parts of Mediator. The jaws and a central joint are implicated in interactions with Pol II and its carboxy-terminal domain, and the joint is required for transcription *in vitro*. The *S. pombe* head module structure leads to a revised model of the *S. cerevisiae* module, reveals a high conservation and flexibility, explains known mutations, and provides the basis for unravelling a central mechanism of gene regulation.

Extending our structural analysis of Mediator¹³, we determined the crystal structure of a bacterially expressed Med6 variant (*S. pombe* (*Sp*) residues 9–180) at 2.7 Å resolution (Supplementary Figs 1, 2a and Supplementary Table 1). The structure revealed a core domain with a five-stranded antiparallel β-sheet, two pairs of α-helices flanking a conserved groove, and a flexible C-terminal helix α5 (Fig. 1a, b and Supplementary Fig. 3). Bacterial expression also provided the head subcomplex Med17C–Med11C–Med22C, which consists of C-terminal regions in Med17, Med11 and Med22 (ref. 12) and constitutes the fixed jaw¹⁰ (see Methods and Supplementary Fig. 2b). The crystal structure at 3.0 Å resolution (Supplementary Table 1) revealed two subdomains we call ‘tooth’ and ‘nose’ (Fig. 1c). The tooth contains Med11C, Med22C and Med17C residues 383–541 and 594–611. It forms a five-stranded β-sheet that is flanked by five helices, three from Med17C and one each from Med11C and Med22C. The nose comprises Med17C residues 542–580 and 612–687 and forms another five-stranded β-sheet with two flanking helices.

We next prepared the entire *S. cerevisiae* (*Sc*) head module by co-expressing its seven subunits in *Escherichia coli* (see Methods and Supplementary Fig. 2c). The purified module (Fig. 2a) supported activator- and promoter-dependent transcription from three different promoters *in vitro* (Fig. 2b and Supplementary Note 1). Crystals of the *Sc* head module were obtained in 2007, but persistently diffracted to low resolution. We therefore prepared the head module from *Sp* (Supplementary Fig. 2d). The purified module (Fig. 2a) did not crystallize,

but crystals were obtained for a variant lacking three non-essential parts, Med6 residues 181–216, Med17 residues 1–76 and Med20 (Fig. 2a and Supplementary Fig. 2e).

We determined the *Sp* head module structure *de novo* by multiple isomorphous replacement with anomalous scattering (see Methods and Supplementary Fig. 4a). Modelling required selenomethionine sequence markers (Supplementary Fig. 4b) and the four available structures of subcomplexes (Supplementary Fig. 5a–e). Diffraction data to 3.4 Å resolution (Supplementary Table 2) resulted in excellent electron density (Supplementary Fig. 4c, d) and a refined structure that has a free *R* factor (*R*_{free}) of 25.8% (Supplementary Table 2) and contains 87% of the residues in the variant.

The *Sp* head module structure reveals an asymmetric multiprotein assembly that resembles the head of a crocodile with one limb (Figs 2c and 3a). The structure reveals all folds and previously unobserved regions, and is partitioned into eight elements (Fig. 3a, b). The neck submodule is formed by Med6, Med8 and parts of Med11, Med17 and Med22, and consists of ‘spine’, ‘shoulder’, ‘arm’ and ‘finger’ elements (Fig. 3a, b and Supplementary Figs 6–8). The spine comprises seven helices, six forming a long bundle (Med8 α5, Med17 α3, Med11N–Med22N (ref. 12)), and a seventh perpendicular helix (Med8 α4). The arm consists of a short four-helix bundle (Med6 α5, Med8 α1/2, Med17 α1), two pairs of helices (Med17 η1/α2, Med8 α2′/α3), and a two-stranded sheet (Med8 β1, Med17 β1). The arm binds the shoulder, which consists of the Med6 core (Fig. 1a). The finger is an exposed C-terminal β-hairpin in Med6 (Fig. 3a, d). The fixed and moveable

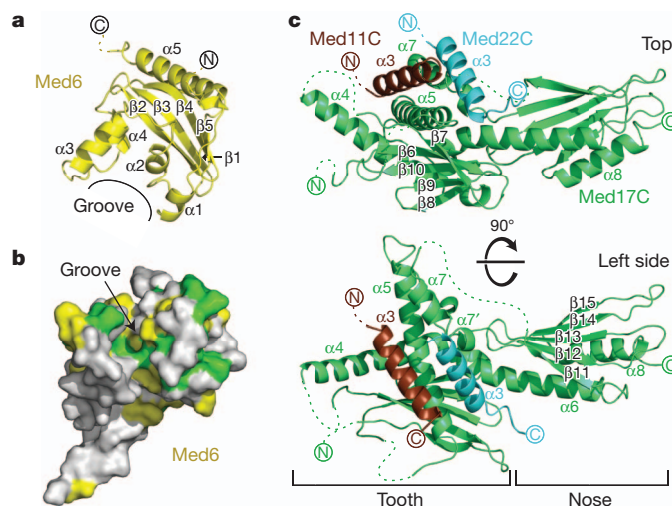


Figure 1 | Structures of Med6 and Med17C–Med11C–Med22C. **a**, Ribbon model of the *Sp* Med6 crystal structure. **b**, Conserved surface groove in Med6. Residues that are invariant or conserved among seven yeast species are in green or yellow, respectively (see Supplementary Fig. 1). **c**, Ribbon model of the *Sc* subcomplex Med17C–Med11C–Med22C crystal structure. ‘Tooth’ and ‘nose’ domains are indicated. Dashed lines indicate disordered regions.

¹Gene Center and Department of Biochemistry, Center for Integrated Protein Science Munich (CIPSM), Ludwig-Maximilians-Universität München, Feodor-Lynen-Straße 25, 81377 Munich, Germany.

[†]Present address: Institute for Molecular Bioscience, University of Queensland, Brisbane St Lucia, Queensland 4072, Australia.

*These authors contributed equally to this work.

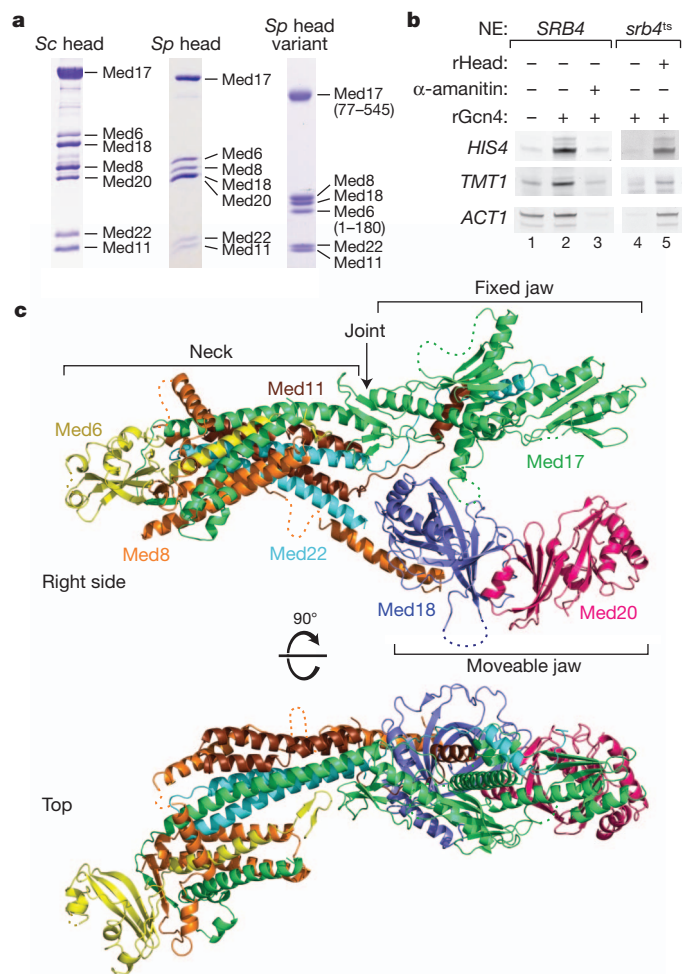


Figure 2 | Structure of *Sp* Mediator head module. **a**, SDS–polyacrylamide gel electrophoresis (PAGE) analysis of purified recombinant head modules of *Sc* (left), full-length *Sp* (middle) and the crystallized *Sp* variant (right). **b**, Recombinant *Sc* head module supports *in vitro* transcription from three different promoter types. Activity of nuclear extracts from *SRB4* and *srb4-138* (*srb4*^{ts}) strains was assessed on promoters *HIS4*, *TMT1* and *ACT1*. Pol II-dependency was confirmed by α -amanitin sensitivity (0.04 mM, lane 3). Addition of 5 pmol recombinant *Sc* head module (lane 5) to inactive *srb4*^{ts} extracts (lane 4) restored activity. **c**, Two views of a ribbon representation of *Sp* Mediator head module structure. Subunits are in different colours. Subunit Med20 is not present in the crystals but was unambiguously modelled and positioned by superimposing the previous *Sc* Med18–Med20 heterodimer structure¹¹.

jaws consist of subcomplexes Med17C–Med11C–Med22C (Figs 1c and 3a) and Med8C–Med18–Med20 (ref. 11), respectively.

Both jaws are connected to the spine. The fixed jaw is connected via the ‘joint’ that contains a four-stranded β -sheet formed by Med17 residues 215–259, and short linkers in Med11 and Med22. The joint may change its structure as it contains many conserved glycine and serine residues. Deletion of the Med17 part of the joint did not cause lethality in yeast, but essentially abolished activated transcription *in vitro* (Supplementary Fig. 9a, b). The moveable jaw binds the spine via Med8C, a helix extending from the Med18–Med20 heterodimer¹⁴ (Fig. 3a, d), and also contacts the joint and tooth. Conservation of the module surface is highest at the shoulder and arm (Fig. 3c and Supplementary Fig. 6), probably because this region binds the middle module^{15,16}. Deletion of the shoulder caused lethality (Supplementary Fig. 9c).

We prepared a homology model for the *Sc* head module based on our structures of the *Sp* module and the *Sc* subcomplexes

Med17C–Med11C–Med22C, Med11N–Med22N (ref. 12) and Med8C–Med18–Med20 (ref. 11). After minor adjustments, the model explained the published electron density and selenium positions¹⁰ (Supplementary Fig. 10a). Compared with the published model¹⁰, four α -helices were swapped, a β -sheet was added in the tooth, and the amino acid register was adjusted for 46% of the residues outside the moveable jaw (Supplementary Fig. 10b). Thus the structures of the *Sp* and *Sc* head modules are well conserved although the structured regions show only 15% sequence identity (Supplementary Fig. 1). *Sc*-specific differences are found in six helices; three are shorter (Med8 α C, Med17 α 5, Med22 α 1), one is longer (Med11 α 1) and two are rotated (Med17 α 7, Med22 α 3) (Supplementary Fig. 5b, d, e). Because of poor electron density¹⁰, the revised *Sc* model lacks the joint, part of the arm, the sheet in the shoulder, the finger and five protein linkers.

Structural comparisons reveal flexibility within the head module. The position of the shoulder changes in different *Sp* module crystals (Supplementary Fig. 11a), owing to a flexible connection to Med6 helix α 5 (Supplementary Figs 3 and 5c). The finger is mobile in the free Med6 structure. Flexibility of the moveable jaw is indicated by high B-factors (Supplementary Fig. 11b and Supplementary Note 2) and electron microscopic results¹⁷. Movement of the nose with respect to the tooth is revealed by comparison of the *Sc* model with the free fixed jaw structure (Supplementary Fig. 11c). Comparison of the *Sp* structure with the *Sc* model indicates that the neck can rotate with respect to the jaws (Supplementary Fig. 11e). The rotation may be accommodated by the joint, which uses conserved hydrophobic residues to anchor the tooth to the spine. Thus the head module contains at least four mobile peripheral elements, the shoulder, finger, moveable jaw and nose.

To explain the phenotypes of known mutations, we mapped 47 sites of 55 *Sc* head module mutations onto the *Sp* structure (Fig. 4, Supplementary Fig. 1 and Supplementary Table 3). Most temperature-sensitive mutations map to protein cores or domain interfaces and apparently influence module stability (Fig. 4a). These include mutant *srb4-138* that contains six Med17 mutations required for temperature sensitivity¹⁸, including mutations Ser226Pro and Phe649Ser that apparently destabilize the arm and nose, respectively. The human MED17 mutation Leu371Pro (residues 504 and 389 in *Sc* and *Sp*, respectively) is associated with infantile cerebral atrophy³ and perturbs MED17 helix α 5 in the tooth. The temperature-sensitive phenotype for Med11 mutations Glu17Lys/Leu24Lys (ref. 12), Leu66Pro (refs 12, 19), and Glu92Ser (refs 12, 19) is also explained by fold destabilizations. The mutation *med6-ts1* and mutations in the *med6-ts6* allele²⁰ may destabilize the shoulder. Mutant *med6-ts2* contains six point mutations²⁰, of which Gln49Leu changes the conserved surface implicated in binding other Mediator parts, Phe125Tyr changes the shoulder–arm interface, and Phe194Leu may destabilize the finger–spine interface.

The structure also explains how mutant phenotypes are suppressed by secondary mutations. The *SRB4-101* mutation Glu286Lys rescues the *med6-ts2* phenotype²⁰, apparently by creating a new salt bridge between Med17 helix α 3 and Med6 helix α 5. The *MED6-101* mutation²¹ may suppress the *srb4-138* phenotype because mutation Asp152Tyr between the shoulder and arm compensates for decreased stability of the neck. The *SRB6-201* (ref. 21) mutation Asn59His may also suppress the *srb4-138* phenotype by stabilizing the neck.

Other genetic data implicate the fixed jaw in interactions with Pol II (Fig. 4b, Supplementary Fig. 1 and Supplementary Table 3). The head module interacts with Pol II near subunit Rpb3, and this contact is required for regulated transcription^{22,23}. Mutations causing co-lethality with the Rpb3 mutation *rpb3-2* map to the fixed jaw (Fig. 4b). Mutants *med17-208* and *med17-257* both contain the mutation Glu669Asp that is located on the nose surface and may abolish the head–polymerase interaction that is weakened by Rpb3 mutation. Mutant *med17-sup1* rescues the phenotype of Rpb3 mutation Ala159Gly (ref. 22) and contains two mutations in the fixed jaw (Fig. 4b). Consistent with an interaction between the head module and the Pol II region around

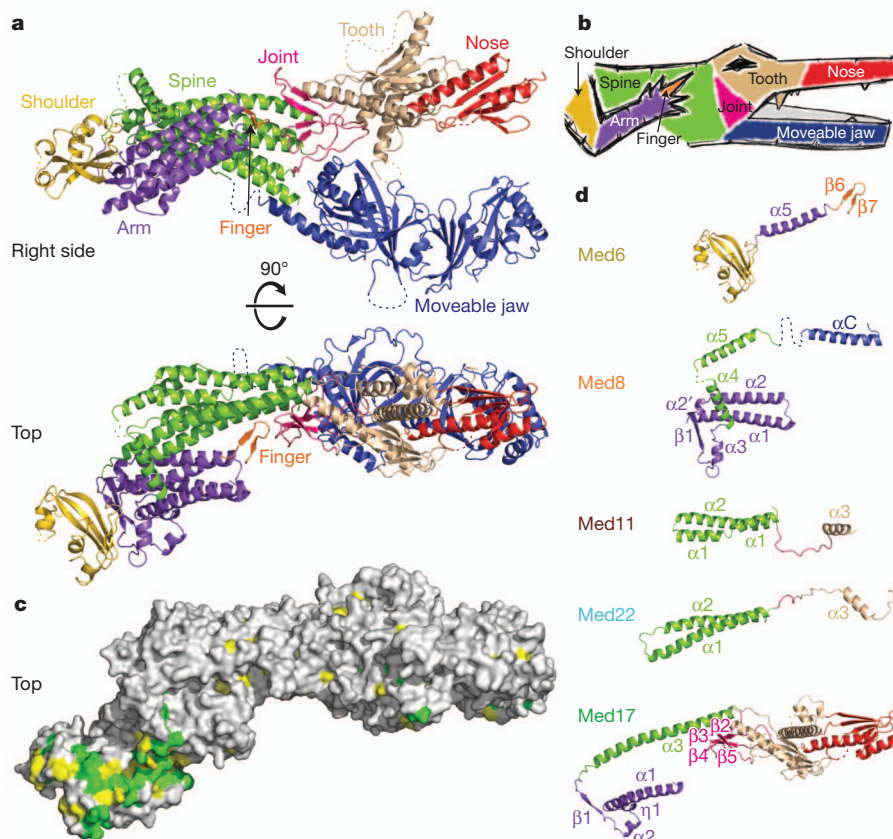
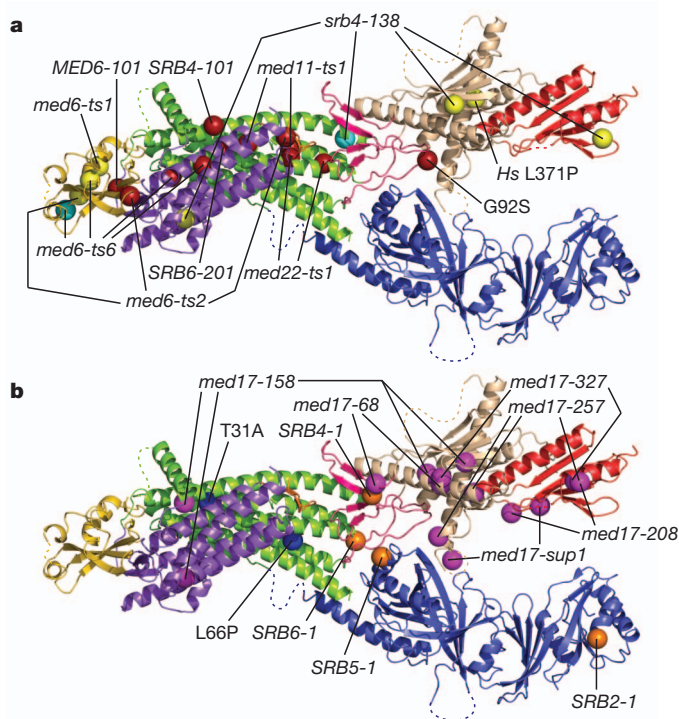


Figure 3 | Structural elements and surface conservation. **a**, Eight structural elements in the Mediator head module. Ribbon model of the *Sp* head module structure is coloured accordingly (see also Supplementary Figs 6–8). **b**, The structure resembles the head of a crocodile with one limb. The eight structural elements received corresponding names, and previously introduced names for the three submodules neck, fixed jaw and moveable jaw¹⁰ are preserved.

c, Surface conservation of the *Sp* Mediator head module. Residues that are invariant or conserved among seven yeast species (Supplementary Fig. 1) are in green or yellow, respectively (see also Supplementary Fig. 6). **d**, Ribbon representation of the five non-globular head subunits (top view) coloured according to structural elements as in **a** (see also Supplementary Figs 7 and 8).



Rpb3, a mutation in the adjacent polymerase dock domain has a cold-sensitive phenotype that is rescued by *SRB* mutations²⁴ in the joint and moveable jaw (Supplementary Fig. 12, *rpb1-14*). Mutants *med17-68*, *med17-158* and *med17-327* all contain mutations that are predicted to destabilize the fixed jaw.

Mediator also interacts with the carboxy-terminal domain (CTD)^{4,25} that extends from the polymerase core near the dock domain (Supplementary Fig. 12). CTD truncation causes a cold-sensitive phenotype that is rescued by mutations^{24,26,27} that map mainly to the joint (Fig. 4b). The mutation *SRB4-1* maps to the joint–tooth interface, whereas the mutation *SRB6-1* may stabilize the spine and adjacent joint. The mutation *SRB5-1* in Med18 helix $\alpha 1$ may strengthen joint interaction with the moveable jaw, whereas mutation *SRB2-1* alters the moveable jaw surface. These results suggest that the joint and moveable

Figure 4 | Head module integrity and interactions. **a**, Location of sites of *Sc* mutations that may influence module stability. Mutations were mapped onto the *Sp* head module structure using alignments shown in Supplementary Fig. 1. C α atoms of mutated residues are shown as spheres. Mutations that may affect intrasubunit stability, intersubunit stability or with uncertain effects are in yellow, dark red or cyan, respectively. **b**, Location of sites of *Sc* mutations that may influence interactions with the Pol II core, CTD and TFIIF. Mutations were mapped onto the *Sp* head module using alignments shown in Supplementary Fig. 1, with C α atoms represented as spheres. Mutations that suppress CTD truncation phenotypes, interact genetically with Rpb3 mutations or lead to a loss of interaction with TFIIF are shown as orange, magenta or blue spheres, respectively. A Med11 mutation in the spine, Thr31Ala, leads to reduced interaction with TFIIF (ref. 19).

jaw are involved in CTD binding and/or that they are required for structural changes in Mediator that enable CTD binding. However, other Mediator regions probably contribute to interactions with the CTD because the head module is insufficient for CTD binding *in vitro*²⁸. The flexibility and extended shape of the head module, which is 170 Å long, 95 Å high and 65 Å wide, may allow for several interactions not only with Pol II but also with transcription factors TFIIB^{6,29}, TFIID^{12,19} and the TATA box-binding protein¹¹.

METHODS SUMMARY

To enable heterologous co-expression of head module subunits and their variants in *E. coli*, more than 300 expression vectors were tested. Expression vectors used in this work are shown in Supplementary Fig. 2. All proteins were expressed in *E. coli* BL21(DE3)RIL cells and purified by nickel affinity, anion exchange and size-exclusion chromatography. Crystals were grown by vapour diffusion at 20 °C. *Sp* Med6 crystals were grown using reservoir solutions containing 100 mM HEPES, pH 7.5, and 400 mM sodium citrate or 16% (v/v) tacsimite for native or selenomethionine-labelled protein, respectively. *Sc* Med17C–Med11C–Med22C crystals were grown in 100 mM MES, pH 6.0, and 4 M ammonium acetate. *Sp* head module crystals were grown in 50 mM MES, pH 6.0, and 1 M ammonium sulphate. Diffraction data were collected under cryo-conditions at the Swiss Light Source (SLS) beamlines X06SA and X06DA. The Med6 and Med17C–Med11C–Med22C structures were solved using selenomethionine labelling and single-wavelength anomalous dispersion phasing. The head module structure was solved by MIRAS using tantalum and ytterbium derivatives and 23 anomalous difference Fourier peaks in a selenomethionine-labelled crystal as sequence markers. Yeast nuclear extract preparation and transcription assays were performed as described^{12,30}.

Full Methods and any associated references are available in the online version of the paper.

Received 10 July; accepted 12 October 2012.

Published online 31 October 2012.

- Conaway, R. C. & Conaway, J. W. Origins and activity of the Mediator complex. *Semin. Cell Dev. Biol.* **22**, 729–734 (2011).
- Chen, W. & Roeder, R. G. Mediator-dependent nuclear receptor function. *Semin. Cell Dev. Biol.* (2011).
- Kaufmann, R. *et al.* Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am. J. Hum. Genet.* **87**, 667–670 (2010).
- Myers, L. C. *et al.* The Med proteins of yeast and their function through the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **12**, 45–54 (1998).
- Asturias, F. J., Jiang, Y. W., Myers, L. C., Gustafsson, C. M. & Kornberg, R. D. Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **283**, 985–987 (1999).
- Kang, J. S. *et al.* The structural and functional organization of the yeast mediator complex. *J. Biol. Chem.* **276**, 42003–42010 (2001).
- Davis, J. A., Takagi, Y., Kornberg, R. & Asturias, F. Structure of the yeast RNA polymerase II holoenzyme: mediator conformation and polymerase interaction. *Mol. Cell* **10**, 409–415 (2002).
- Liu, Y., Ranish, J. A., Aebersold, R. & Hahn, S. Yeast nuclear extract contains two major forms of RNA polymerase II mediator complexes. *J. Biol. Chem.* **276**, 7169–7175 (2001).
- Bourbon, H. M. Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res.* **36**, 3993–4008 (2008).
- Imasaki, T. *et al.* Architecture of the Mediator head module. *Nature* **475**, 240–243 (2011).
- Larivière, L. *et al.* Structure and TBP binding of the Mediator head subcomplex Med8–Med18–Med20. *Nature Struct. Mol. Biol.* **13**, 895–901 (2006).
- Seizl, M., Larivière, L., Pfaffeneder, T., Wenzek, L. & Cramer, P. Mediator head subcomplex Med11/22 contains a common helix bundle building block with a specific function in transcription initiation complex stabilization. *Nucleic Acids Res.* **39**, 6291–6304 (2011).
- Larivière, L., Seizl, M. & Cramer, P. A structural perspective on Mediator function. *Curr. Opin. Cell Biol.* (2012).
- Larivière, L. *et al.* Structure-system correlation identifies a gene regulatory Mediator submodule. *Genes Dev.* **22**, 872–877 (2008).
- Baumli, S., Hoepfner, S. & Cramer, P. A conserved mediator hinge revealed in the structure of the MED7–MED21 (Med7–Srb7) heterodimer. *J. Biol. Chem.* **280**, 18171–18178 (2005).
- Guglielmi, B. *et al.* A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res.* **32**, 5379–5391 (2004).
- Cai, G. *et al.* Mediator head module structure and functional interactions. *Nature Struct. Mol. Biol.* **17**, 273–279 (2010).
- Takagi, Y. & Kornberg, R. D. Mediator as a general transcription factor. *J. Biol. Chem.* **281**, 80–89 (2006).
- Esnault, C. *et al.* Mediator-dependent recruitment of TFIID modules in preinitiation complex. *Mol. Cell* **31**, 337–346 (2008).
- Lee, Y. C. & Kim, Y. J. Requirement for a functional interaction between mediator components Med6 and Srb4 in RNA polymerase II transcription. *Mol. Cell Biol.* **18**, 5364–5370 (1998).
- Lee, T. I. *et al.* Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. *Mol. Cell Biol.* **18**, 4455–4462 (1998).
- Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C. & Werner, M. Direct interaction of RNA polymerase II and mediator required for transcription *in vivo*. *Science* **331**, 1451–1454 (2011).
- Tan, Q., Linask, K. L., Ebright, R. H. & Woychik, N. A. Activation mutants in yeast RNA polymerase II subunit RPB3 provide evidence for a structurally conserved surface required for activation in eukaryotes and bacteria. *Genes Dev.* **14**, 339–348 (2000).
- Thompson, C. M., Koleske, A. J., Chao, D. M. & Young, R. A. A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**, 1361–1375 (1993).
- Näär, A. M., Taatjes, D. J., Zhai, W., Nogales, E. & Tjian, R. Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev.* **16**, 1339–1344 (2002).
- Nonet, M. L. & Young, R. A. Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics* **123**, 715–724 (1989).
- Koleske, A. J., Buratowski, S., Nonet, M. & Young, R. A. A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. *Cell* **69**, 883–894 (1992).
- Takagi, Y. *et al.* Head module control of mediator interactions. *Mol. Cell* **23**, 355–364 (2006).
- Baek, H. J., Kang, Y. K. & Roeder, R. G. Human Mediator enhances basal transcription by facilitating recruitment of transcription factor IIB during preinitiation complex assembly. *J. Biol. Chem.* **281**, 15172–15181 (2006).
- Ranish, J. A., Yudkovsky, N. & Hahn, S. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* **13**, 49–63 (1999).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. Baumli, A. C. Cheung, A. Imhof and C. Schmidt for help. We acknowledge the crystallization facility at the Max Planck Institute of Biochemistry, Martinsried. Diffraction data were collected at the Swiss Light Source, Villigen, Switzerland. M.S. was supported by a Boehringer Ingelheim fellowship and the Elite Network of Bavaria. P.C. was supported by the Deutsche Forschungsgemeinschaft, SFB646, TR5, GraKo1721, SFB960, CIPSM, NIM, an Advanced Grant of the European Research Council, the LMUinnovativ project Bioimaging Network, the Vallee Foundation, and the Jung-Stiftung.

Author Contributions L.L. designed crystallization constructs, established expression and purification strategies for head modules, solved the Med6 and Med17C–Med11C–Med22C structures, carried out crystallographic data analysis for *Sp* head module structure determination, built the *Sc* head module model, generated yeast strains, and performed *in vivo* studies. C.P. optimized *Sp* head module purification, purified and crystallized the *Sp* head module variants, and prepared heavy metal derivatives. L.L. and C.P. collected diffraction data for the head module and carried out model building and structural analysis. M.S. contributed to establishing expression and purification strategies and carried out transcription assays. L.W. provided technical help. F.K. optimized transcription assays. P.C. initiated and supervised the project. P.C., L.L., C.P. and M.S. prepared the manuscript.

Author Information Atomic coordinate files and structure factors were deposited in the Protein Data Bank under accession codes 4H61 (*Sp* Med6), 4H62 (*Sc* Med17C–Med11C–Med22C) and 4H63 (*Sp* head module). See Supplementary Data for the coordinate file for the revised *Sc* head module model. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.L. (larivier@genzentrum.lmu.de) or P.C. (cramer@genzentrum.lmu.de).

METHODS

Vectors and sequences. To enable heterologous co-expression of head module subunits and their variants in *E. coli*, more than 300 expression vectors were tested. Vectors used in this work are shown in Supplementary Fig. 2. Vectors containing a single T7 promoter are based on pET21, pET24 or pET28 vectors (Novagen). Vectors containing two T7 promoters are based on Duet vectors (Novagen). The head co-expression vector that contains three T7 promoters was constructed by concatenating two Duet vectors. Open reading frames (ORFs) were cloned sequentially into vectors. For polycistronic constructs, ribosomal-binding sites were introduced as described¹⁵. Sequences of all constructs are available on request. All proteins were expressed in *E. coli* BL21(DE3)RIL (Stratagene). Cells were grown in LB medium at 37 °C to an attenuation (*D*) of 0.5 at 600 nm. Expression was induced with 0.5 mM isopropyl- β -D-thiogalactoside (IPTG) for 16 h at 18 °C. Sequence alignments were performed with MUSCLE³¹ and manually adjusted based on the presented structures. Graphical representation in Supplementary Fig. 1 was generated with ESPript³².

Sp Med6 crystal structure determination. For expression of *Sp* Med6 (residues 9–180), cells were transformed with the vector depicted in Supplementary Fig. 2a and the protein was purified as described¹² except that a HiLoad 16/600 Superdex 200 pg (GE Healthcare) was used for size exclusion chromatography. Crystals were grown at 20 °C in hanging drops over reservoirs containing 100 mM HEPES, pH 7.5, and 400 mM sodium citrate. For crystallization of selenomethionine-labelled³³ protein, sodium citrate was replaced by 16% (v/v) tacsimate. Crystals were collected, cryo-protected by gradually adding glycerol to a final concentration of 30% (v/v), and flash-frozen in liquid nitrogen. Native and selenomethionine single-wavelength anomalous dispersion (SAD) diffraction data were collected at 100 K on a PILATUS 6M detector at the X06SA beamline at Swiss Light Source (SLS), Villigen, Switzerland. Data were processed with XDS and XSCALE³⁴. Phenix.autosol³⁵ was used to detect selenium sites, phase the structure, perform density modification, and for initial model building. The model was manually adjusted using COOT³⁶ and refined with Phenix.refine³⁵, including high-resolution data³⁷ because the resulting final model had a better R_{free} value and stereochemistry. The refined structure has an R_{free} value of 21.7% and shows very good stereochemistry (Supplementary Table 1). Ninety-six per cent of the residues fall in favoured regions of the Ramachandran plot and none of them is in disallowed regions³⁸.

Sc Med17C–Med11C–Med22C crystal structure determination. *Sc* Med17C–Med11C–Med22C was expressed in cells co-transformed with the two vectors depicted in Supplementary Fig. 2b and was purified as *Sp* Med6 above. For the Med11C–Med22C fusion construct, a linker with sequence GAGSGAGSG was inserted between the C terminus of Med11 and residue 96 of Med22. This covalent linker was essential for complex stability. Crystals were grown at 20 °C in hanging drops over reservoirs containing 100 mM MES, pH 6.0, and 4 M ammonium acetate. Crystals were collected, transferred to a solution containing 100 mM MES, pH 6.0 and 5 M ammonium acetate, and flash-frozen in liquid nitrogen. Selenomethionine labelling and diffraction data collection were as above for *Sp* Med6. Data were processed with XDS and XSCALE³⁴. SHELCX/D/E³⁹ was used to detect selenium sites, phase the structure, and perform density modification. The resulting electron density map allowed for building of most of the model with COOT³⁶. The model was refined with Phenix.refine³⁵, including high-resolution data³⁷ because the resulting final model had a better R_{free} value and stereochemistry. The refined structure has a R_{free} value of 23.3% and shows very good stereochemistry (Supplementary Table 1). Ninety-six of the residues fall in favoured regions of the Ramachandran plot and none of them is in disallowed regions³⁸.

Preparation of recombinant *Sc* head module. *Sc* head module was expressed in cells transformed with the single vector depicted in Supplementary Fig. 2c. Cells were lysed by sonication in buffer A (50 mM HEPES-potassium hydroxide, pH 7.5, 400 mM potassium chloride, 10% (v/v) glycerol, 5 mM dithiothreitol (DTT)) containing protease inhibitors¹². After centrifugation, the supernatant was loaded onto a 1 ml HiTrap column (GE Healthcare) equilibrated in buffer B (50 mM HEPES-potassium hydroxide, pH 7.5, 500 mM potassium acetate, 10% (v/v) glycerol, 50 mM imidazole, 5 mM DTT). The complex was eluted with a linear gradient from 50 mM to 300 mM imidazole in buffer B. The *Sc* head module was further purified by anion exchange chromatography with a 1 ml HiTrap Q HP column (GE Healthcare). The column was equilibrated in buffer C (50 mM HEPES-potassium hydroxide, pH 7.5, 150 mM potassium acetate, 10% (v/v) glycerol, 2 mM DTT), and proteins were eluted with a linear gradient from 150 mM to 1.25 M potassium acetate in buffer C. Fractions containing head module were applied to a HiLoad 16/600 Superdex 200 pg (GE healthcare) size exclusion column equilibrated in buffer D (20 mM HEPES-potassium hydroxide, pH 7.5, 150 mM potassium acetate, 10% (v/v) glycerol, 2 mM DTT). The *Sc* head module was concentrated to 5 mg ml^{−1}, flash-frozen, and stored at −80 °C. The identity of the polypeptides was confirmed by mass spectrometry. Static light scattering showed that the complex contains one copy of each subunit (not shown). The

Sc head module variant lacking the Med17 part of the joint (Med17_{Δ321–369}) was expressed and purified the same way.

Yeast strains and functional assays. Plasmids pRS316-SRB4 and pRS316-MED6 were generated by cloning the respective ORF plus 500 base pairs (bp) upstream and 300 bp downstream sequence into pRS316 (ATCC; *URA3* marker). Plasmids pRS315-SRB4, pRS315-srb4⁴⁵, pRS315-med17_{Δ321–369}, pRS315-MED6, pRS315-med6_{1–225} and pRS315-med6_{163–295} were generated by cloning the respective wild-type or mutant ORF plus 500 bp upstream and 300 bp downstream sequence into pRS315 (ATCC; *LEU2* marker). The heterozygous *MED17/med17Δ* and *MED6/med6Δ* *Sc* yeast strains were obtained from Euroscarf and transformed with pRS316-SRB4 and pRS316-MED6, respectively. Diploids were sporulated, tetrads were dissected and analysed, and a shuffle strain was selected. To assess functionality of mutants, pRS315 constructs were transformed into the respective shuffle strain. Equal amounts of freshly grown yeast cells in SC (−*URA* −*Leu*) medium were resuspended in water and tenfold dilutions were spotted on 5-fluoroorotic acid (5-FOA) and SC (−*URA* −*Leu*) plates. Isogenic *SRB4* and *srb4*⁴⁸ strains used for nuclear extract preparation were obtained by transforming the *MED17* shuffle strain with plasmids pRS315-SRB4 or pRS315-srb4⁴⁵ and streaking transformants twice on 5-FOA plates. Nuclear extracts were prepared from 3 l of yeast cultures as described^{12,30}. *In vitro* transcription and analysis by primer extension were performed as described¹². Primer extension was done using the same 5'-Cy5-labelled oligonucleotide (5'-TTCACCACTGAGACGGGCAAC-3') for all promoters tested. For activated transcription we added 200 ng of recombinant full-length Gcn4 that was purified as described⁴⁰. *HIS4* and *TMT1* template plasmids were as described⁴⁰. The *ACT1* template plasmid was generated by inserting the *ACT1* promoter sequence (425 bp upstream to 24 bp downstream of the start codon) in pBluescript KS+ with HindIII and BamHI.

Preparation of recombinant *Sp* head module. The complete *Sp* head module was expressed in cells co-transformed with the three plasmids depicted in Supplementary Fig. 2d. The crystallized *Sp* head module was expressed in cells co-transformed with the three plasmids shown in Supplementary Fig. 2e. Cells were lysed by sonication in buffer E (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, 5 mM DTT) containing protease inhibitors¹². After centrifugation, the supernatant was precipitated with 35% (v/v) saturated ammonium sulphate, and pellets were resuspended in buffer E and loaded on a 2 ml Ni-NTA agarose beads column (QIAGEN) equilibrated in buffer E. The flow-through was reloaded on another 2 ml Ni-NTA column and both columns were washed with buffer E containing increasing concentration of imidazole (0, 10 and 20 mM). The complex was eluted from both columns with buffer E containing 300 mM imidazole. The His-tag was cleaved overnight with 0.7 U thrombin (Sigma-Aldrich) per mg of protein while dialysing against buffer F (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM DTT). The complex was further purified by anion exchange chromatography using a 1 ml HiTrap Q HP column (GE Healthcare) equilibrated in buffer G (20 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10% (v/v) glycerol, 2 mM DTT) and eluted with a linear gradient from 50 mM to 500 mM NaCl in buffer G over 70 column volumes. Fractions containing the complex were collected and applied to a HiLoad 16/600 Superdex 200 pg (GE Healthcare) size exclusion column equilibrated in buffer F. The *Sp* head module was concentrated to 6 mg ml^{−1}, flash-frozen and stored at −80 °C. The identity of the polypeptides was confirmed by mass spectrometry.

Crystal structure determination of *Sp* Mediator head module. Crystals of the native complex were grown at 3.5 mg ml^{−1} *Sp* head module at 20 °C in hanging drops over reservoirs containing 50 mM MES, pH 6.0, and 1 M ammonium sulphate. Crystals were collected and transferred in a stepwise manner to the final cryo-solution (50 mM MES, pH 6.0, 100% saturated lithium sulphate) and flash-frozen in liquid nitrogen. For structure solution native crystals were derivatized with Ta₆Br₁₂ (refs 41–43) (Proteros biostructures) and Yb-DTPA-BMA⁴⁴ (NatX-ray). Ta₆Br₁₂ was added directly to the crystallization drop at 2 mM for 1 h. Yb-DTPA-BMA was added to the final cryo-solution at 100 mM for 10 min and back-soaked 10 s before freezing. The *Sp* head module was further labelled with selenomethionine as described³³. Crystals of the labelled protein were grown at 2 mg ml^{−1} *Sp* head module at 20 °C in hanging drops over a reservoir composed of 50 mM MES, pH 6.0, 0.95 M ammonium sulphate. Crystals were measured at 100 K at SLS, Villigen, Switzerland. Diffraction data were collected on a PILATUS 6M detector at the X06SA beamline (native crystals and Ta₆Br₁₂ and Yb-DTPA-BMA derivatives) or on a PILATUS 2M detector at the X06DA beamline (selenomethionine-labelled crystals). Data were processed with XDS and XSCALE³⁴. The HySS submodule in Phenix³⁵ identified four heavy-atom sites in the Ta₆Br₁₂ derivative and a single heavy-atom site in the Yb-DTPA-BMA derivative. These sites were used as input for MIRAS phasing with autoSHARP⁴⁵. Density modification was performed with Resolve⁴⁶ and yielded an interpretable electron density map. In parallel, an anomalous difference Fourier map was calculated with data from selenomethionine-labelled crystals and with experimental MIRAS phases. The experimental map, in combination with the selenomethionine

sequence markers, allowed the unambiguous manual fitting of the amino-terminal part of Med6 (residues 10–117) and of the Med8C–Med18 subcomplex (PDB code 3COT). Guided by our knowledge of the folds and topologies of the Med11N–Med22N and Med17C–Med11C–Med22C subcomplexes, we could build a poly-alanine model corresponding to all remaining secondary structure elements with COOT³⁶. MIRAS phases were then combined with phases from this initial model and density was modified using the AutoBuild routine in Phenix³⁵. This led to an improved electron density map that allowed for an extension of the initial model. This process was iterated until no further improvement of the map was observed. Then, 23 selenium sites were identified and used as input for SAD phasing with the AutoSol routine in Phenix³⁵. The new phases were combined with phases from the partial model. After density modification, an electron density map of excellent quality was obtained, in which many side chains and most of the linkers between secondary structure elements were visible. After rebuilding, the model was refined with autoBUSTER⁴⁷. The resulting $2F_o - F_c$ electron density map allowed further improvement of the model. After several rounds of rebuilding and refinement with autoBUSTER and Phenix.refine³⁵ using individual B-factor and TLS refinement, we obtained the final model, which shows good stereochemistry and an R_{free} value of 25.8% (Supplementary Table 2). Including high-resolution data³⁷ resulted in a refined structure with better R_{free} values and stereochemistry. More than 95% and 99% of the residues fall in favoured and allowed regions of the Ramachandran plot, respectively³⁸.

Modelling of the Sc Mediator head module. A hybrid model was built using our structure of Sc Med17C–Med11C–Med22C and the structure of Sc Med8C–Med18–Med20 (ref. 11) as models for the fixed and moveable jaws, respectively. For the neck, our Sp head module was used, replacing Med11N and Med22 helix $\alpha 1$ with the corresponding Sc elements in the Sc Med11N–Med22N structure¹². Other amino acids of the neck were replaced with their Sc counterparts, based on sequence alignments. Hydrophobic character of the amino acid situated in the core in the resulting neck model was systematically checked. The obtained models for the two jaws, the spine and the shoulder were separately fitted onto the corresponding elements in the published architectural model of the Sc head module¹⁰ (PDB code 3RJ1), using secondary structure matching in COOT³⁶. This resulted in a model that explained most of the published electron density. Regions of the model that lacked convincing density were removed, as well as side-chain atoms. The resulting model was adjusted by rigid-body refinement with Phenix.refine³⁵,

using the six structural elements present in the model (partial shoulder, partial arm, spine, tooth, nose and moveable jaw) as separate rigid-body groups. For figure preparation and structure interpretation, Sp Med20 was modelled with MODELLER⁴⁸.

31. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
32. Gouet, P., Courcelle, E., Stuart, D. I. & Metz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308 (1999).
33. Budisa, N. *et al.* High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eu. J. Biochem.* **230**, 788–796 (1995).
34. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
35. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
36. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
37. Karplus, P. A. & Diederichs, K. Linking crystallographic model and data quality. *Science* **336**, 1030–1033 (2012).
38. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
39. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122 (2008).
40. Seizl, M. *et al.* A conserved GA element in TATA-less RNA polymerase II promoters. *PLoS ONE* **6**, e27595 (2011).
41. Schneider, G. & Lindqvist, Y. Ta6Br14 is a useful cluster compound for isomorphous replacement in protein crystallography. *Acta Crystallogr. D* **50**, 186–191 (1994).
42. Knäblein, J. *et al.* Ta₆Br₁₂²⁺, a tool for phase determination of large biological assemblies by X-ray crystallography. *J. Mol. Biol.* **270**, 1–7 (1997).
43. Cramer, P. *et al.* Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640–649 (2000).
44. Girard, E., Stelter, M., Vicat, J. & Kahn, R. A new class of lanthanide complexes to obtain high-phasing-power heavy-atom derivatives for macromolecular crystallography. *Acta Crystallogr. D* **59**, 1914–1922 (2003).
45. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230 (2007).
46. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
47. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D Biol* **60**, 2210–2221 (2004).
48. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **Chapter 5**, 5.6.1 (2006).

Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning

Matthew T. C. Brown^{1*}, Kelly R. Tan^{1*}, Eoin C. O'Connor^{1*}, Irina Nikonenko¹, Dominique Muller¹ & Christian Lüscher^{1,2}

The ventral tegmental area (VTA) and nucleus accumbens (NAc) are essential for learning about environmental stimuli associated with motivationally relevant outcomes. The task of signalling such events, both rewarding and aversive, from the VTA to the NAc has largely been ascribed to dopamine neurons^{1–3}. The VTA also contains GABA (γ -aminobutyric acid)-releasing neurons, which provide local inhibition^{4,5} and also project to the NAc^{6,7}. However, the cellular targets and functional importance of this long-range inhibitory projection have not been ascertained. Here we show that GABA-releasing neurons of the VTA that project to the NAc (VTA GABA projection neurons) inhibit accumbal cholinergic interneurons (CINs) to enhance stimulus–outcome learning. Combining optogenetics with structural imaging and electrophysiology, we found that VTA GABA projection neurons selectively target NAc CINs, forming multiple symmetrical synaptic contacts that generated inhibitory postsynaptic currents. This is remarkable considering that CINs represent a very small population of all accumbal neurons, and provide the primary source of cholinergic tone in the NAc. Brief activation of this projection was sufficient to halt the spontaneous activity of NAc CINs, resembling the pause recorded in animals learning stimulus–outcome associations^{8–12}. Indeed, we found that forcing CINs to pause in behaving mice enhanced discrimination of a motivationally important stimulus that had been associated with an aversive outcome. Our results demonstrate that VTA GABA projection neurons, through their selective targeting of accumbal CINs, provide a novel route through which the VTA communicates saliency to the NAc. VTA GABA projection neurons thus emerge as orchestrators of dopaminergic and cholinergic modulation in the NAc.

We took advantage of optogenetic projection targeting to visualize and selectively activate VTA GABA projections neurons (GPNs) to the NAc, which represent about 25% of all VTA GABA cells⁷. We stereotactically injected GAD-Cre mice (65-kDa isoform of the *Gad2* locus^{4,13}; also called *Gad2*-CreER) with adeno-associated virus (AAV) containing floxed ChR2-eYFP into the VTA and prepared sagittal slices of the NAc 3 weeks later (Fig. 1a). Enhanced YFP expression allowed the visualization of axonal projections (Fig. 1b). At higher magnification, these revealed corkscrew-like terminal arborizations that were sparsely distributed throughout the NAc (Fig. 1c). When such images were merged with immunohistochemical identification of choline acetyltransferase (ChAT) to label CINs in fluorescence and light microscopy, the axons of the VTA GPNs closely apposed the soma and proximal dendrites of CINs (Fig. 1c, d). Parvalbumin-positive neurons in the NAc, a subclass of GABA interneurons, did not show any axonal appositions (Supplementary Fig. 1). To examine whether the appositions onto CINs gave rise to synaptic contacts, we next prepared slices for electron microscopy to perform an ultrastructural characterization of the appositions. CINs were identified by immunoreactivity for ChAT

revealed by diaminobenzidine (DAB) using a pre-embedding immunoperoxidase reaction, as well as by distinctive invaginations of their nuclei. VTA GABA axons (DAB revelation of anti-GFP antibodies) surrounding the CINs gave rise to multiple symmetrical synaptic contacts typical of GABA synapses (Fig. 1e). Conversely, VTA GABA axons apposed to medium spiny neurons (MSNs), the predominant cell type in the NAc, almost never formed synaptic contacts (Fig. 1f). In fact, we found 8 synapses with eYFP-stained synaptic boutons on 132 cross-sections or approximately 6,300 μm of analysed membrane of CINs and

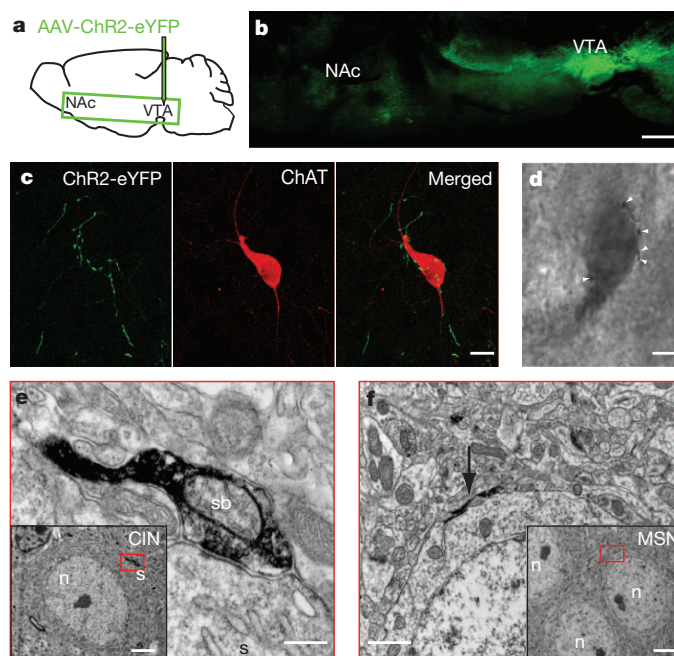


Figure 1 | Long-range GABA projections from the VTA form synaptic contacts selectively with CINs in the NAc. **a**, Schematic diagram showing the injection site and region (green rectangle) used for the sagittal section. **b**, Sagittal section of a GAD-Cre⁺ mouse brain showing eYFP⁺ axons coursing from the VTA to the NAc. Note fibre bundle projecting elsewhere, probably to the prefrontal cortex⁷. **c**, Confocal images of an example NAc CIN immunoreactive for choline acetyltransferase (ChAT) (red) and eYFP⁺ axons (green). **d**, Light microscopy image of pre-embedding immunoperoxidase double labelling of ChAT⁺ cell body and eYFP⁺ axons (arrow heads). **e**, Electron microscopy image of a synaptic contact formed between an eYFP⁺ synaptic bouton (sb) and a CIN. The CIN exhibited a lightly DAB-labelled soma (s, inset), and possessed a typical indented nucleus (n, inset). **f**, An example image of an eYFP⁺ axon apposing (arrow), but not forming a synapse with, an MSN (smaller cell body, round nucleus typical of MSNs with clear cytoplasm (inset)). Scale bars: **b**, 500 μm ; **c**, **d**, 20 μm ; **e**, 0.3 μm and 3 μm (inset); **f**, 1 μm and 3.5 μm (inset).

¹Department of Basic Neurosciences, Medical Faculty, University of Geneva, CH-1211 Geneva, Switzerland. ²Clinic of Neurology, Department of Clinical Neurosciences, Geneva University Hospital, CH-1211 Geneva, Switzerland.

*These authors contributed equally to this work.

only one such synaptic contact on 1,100 cell cross-sections or more than 35,000 μm of MSN membrane (Mann–Whitney U -test, $P < 0.0001$), indicating a selective targeting of NAc CINs by VTA GPNs.

To examine transmission from VTA GPNs onto accumbal neurons further, we performed whole-cell recordings in CINs, MSNs and parvalbumin interneurons while activating VTA GPN axons (Fig. 2a). CINs and MSNs could readily be distinguished by their distinct appearance and electrophysiological properties (Fig. 2b–d; see also refs 14, 15). For the identification of the parvalbumin interneurons we crossed a line that expresses the Cre element in parvalbumin cells with a line expressing a floxed tomato construct (fTom^{PVALBcre}). A brief

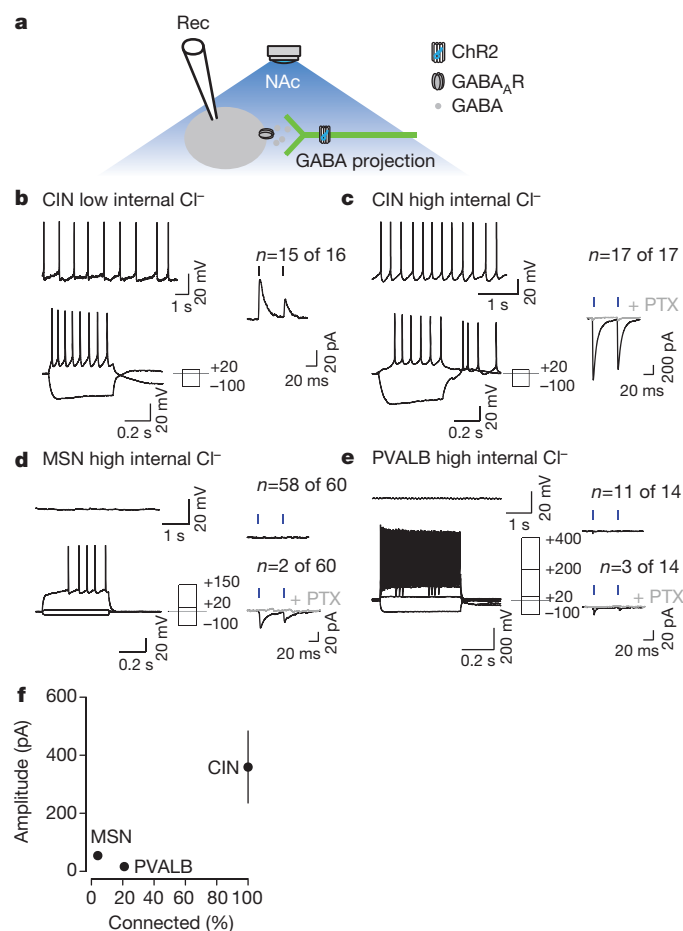


Figure 2 | Inhibitory currents can be selectively elicited in CINs by activating VTA GABA axons in the NAc. **a**, Schematic showing the preparation used to obtain *in vitro* whole-cell patch-clamp recordings of NAc neurons. **b**, Example traces of whole-cell current-clamp recordings of CIN spontaneous activity (left, top panel) or responses to current steps (left, bottom panel) with low internal chloride concentration. An example trace (right) of voltage-clamp recordings of a CIN during a paired-pulse light-flash protocol (4-ms pulses, 20 Hz; representing $n = 15$ of 16) is shown. **c**, Same as in **b** but with high internal chloride concentration ($n = 17$ of 17). IPSCs were blocked by picrotoxin. **d**, Same as in **c** but for an MSN. Most MSNs ($n = 58$ of 60) exhibited no response to light flashes (right, top panel; example trace). In a minority of cases a picrotoxin-sensitive current was observed (right, bottom panel; example trace, $n = 2$ of 60) that was smaller than that observed in CINs. **e**, Same as in **c** and **d** but for identified parvalbumin interneurons. NAc slices were prepared from non-floxed ChR2-infected fTom^{PVALBcre} mice in the VTA. Most parvalbumin interneurons ($n = 11$ of 14) exhibited no IPSCs in response to blue-light stimulation and in a few cases ($n = 3$ of 14), a picrotoxin-sensitive current of much smaller amplitude than in CINs was recorded. Note that in this preparation IPSCs were still selectively recorded in CINs (data not shown). **f**, Graph showing average amplitude (\pm s.e.m.) of IPSCs recorded in CINs, MSNs and parvalbumin interneurons as a function of the percentage under the inhibitory synaptic control of VTA GPNs.

flash of blue light yielded robust inhibitory postsynaptic currents (IPSCs) in 100% of CINs; these were abolished by the GABA_A receptor blocker picrotoxin (Fig. 2c, right panel). In contrast, recording from neighbouring MSNs and parvalbumin interneurons with the same stimulation protocol failed to elicit a response in 96.6% and 80% of cases, respectively, and yielded only small IPSCs in responding neurons (Fig. 2d, e). On average, current amplitudes recorded from CINs were markedly larger than those currents recorded from MSNs and parvalbumin interneurons, if they showed any at all (CINs: 360 ± 124 pA, $n = 17$ of 17 neurons; MSNs: 52 ± 12 pA, $n = 2$ of 60 neurons; parvalbumin interneurons: 20 ± 3 pA, $n = 3$ of 14 neurons; Fig. 2f). The combination of structural imaging as well as electrophysiological evidence *in vitro* demonstrates that VTA GPNs selectively target CINs of the NAc, which is remarkable considering that CINs account for less than 1% of the total number of cells in this region¹⁶.

The connectivity observed here is intriguing because it correlates with the observed *in vivo* responses of CINs and VTA GABA cells to salient events. That is, despite heterogeneity in VTA GABA neuron responses, many are phasically activated¹⁷, whereas striatal CINs pause in response to motivationally important outcomes and associated stimuli^{8–12}, irrespective of their valence. Although the firing pattern of CINs in the NAc has not been well described in animals learning discrete stimulus–outcome associations, our next series of experiments addressed the question of whether activation of VTA GPNs is sufficient to induce a pause in the activity of NAc CINs. We first observed that, in the slice preparation, CINs exhibited tonic firing with an average spontaneous discharge of about 1 Hz. Optogenetic activation of VTA GPN axons was sufficient to transiently silence tonic activity *in vitro*, which was followed by rebound excitation (Fig. 3a). The pause–rebound was abolished in the presence of the GABA_A receptor blocker bicuculline but remained unaffected by either dopamine or acetylcholine (ACh) receptor antagonists (Fig. 3a and Supplementary Fig. 2). We then aimed an optical fibre at the VTA of anaesthetized mice and delivered a light stimulation protocol that was efficient in driving action potentials in VTA GABA neurons *in vivo* (Supplementary Fig. 3) and mimicked physiological activity of these cells (for example, in response to foot shock⁴). When recording from NAc CINs during this same protocol (Fig. 3b), we found that blue light was sufficient to silence spontaneous CIN activity for the entirety of the stimulation (Fig. 3c, e). Again, this pause was followed by rebound excitation, similar to the CIN response observed in the dorsal striatum of behaving monkeys⁹. To verify the accuracy of the electrophysiological criteria used to identify CINs *in vivo* (Supplementary Fig. 4¹⁸), we labelled a number of these recorded neurons with the juxtacellular method to later confirm ChAT expression (Fig. 3d). Together, these data demonstrate that VTA GPN activation is sufficient to pause spontaneous firing of NAc CINs through GABA_A receptor-mediated inhibition.

In correlative studies using non-human primates a ‘conditioned pause response’ of CINs in the dorsal striatum develops as animals learn about a stimulus (for example, a light or tone) that predicts an important outcome (for example, a punishment or a reward^{8–12}). The appearance of the conditioned pause requires several stimulus–outcome pairings and follows a time course that mirrors the occurrence of stimulus-guided anticipatory responses (for example, licking at the spout where a liquid reward will be delivered¹⁹). Given the role of the ventral striatum in associative learning¹, we predicted that activation of VTA GPNs to force a pause response in NAc CINs from the very first stimulus–outcome pairing would facilitate learning about the stimulus–outcome association.

To test this we combined a discrete cue and contextual conditioning task with optogenetically generated pause responses of NAc CINs (Fig. 4a) and compared learning between GAD-Cre⁺ and GAD-Cre[−] mice, both of which were injected in the VTA with AAV-ChR2-eYFP. Mice were first trained across two successive days to associate an auditory cue (conditioned stimulus, CS+; for example,

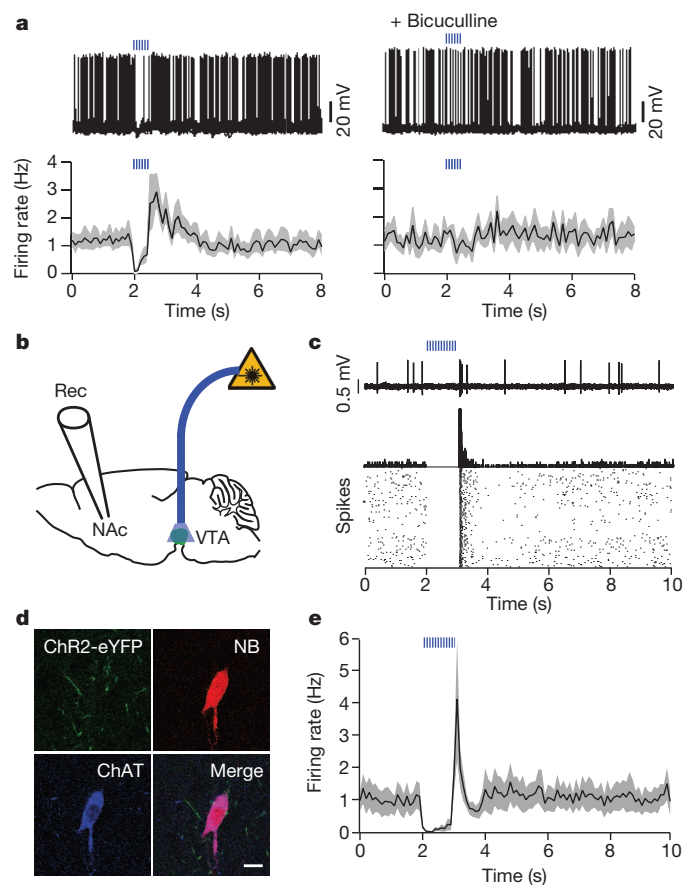


Figure 3 | Pausing NAc CINs through selective VTA GABA projection neuron activation. **a**, Twenty superimposed traces of an example CIN recording *in vitro* in whole-cell current-clamp configuration (top left) shows the blocking of spontaneous action potential firing by VTA GABA axon activation (4-ms pulses, 20 Hz, 0.5 s). This inhibition was often followed by a rebound excitation, as seen in the average response of all recorded neurons (bottom left, $n = 14$). Bath application of 10 μ M bicuculline abolished this inhibition, as shown in the example cell (top right) and the averaged neuronal response (bottom right, $n = 6$). **b**, Schematic showing implantation of a fibre

optic through a guide cannula in the VTA and *in vivo* recordings in the NAc. **c**, Single-unit recording of a neuron inhibited by the light-flash protocol (4-ms pulse, 20 Hz, 1 s) evident in the single sweep (top), the summed peri-stimulus time histogram (PSTH) (middle, 5-ms bins) and the raster plot of all sweeps (bottom, 5-ms bins). **d**, This neuron was labelled with the juxtacellular method and later confirmed as cholinergic using immunohistochemical processing. **e**, Average responses (mean \pm s.e.m., 100-ms bins) of all recorded NAc CINs, some of which were labelled and later confirmed as CINs ($n = 4$ of 11; see also Supplementary Fig. 4).

white noise) with an aversive outcome (a brief foot shock). Mice were also trained with a second stimulus (CS $^-$; for example, pure tone), but had no associated outcome. A light guide aimed at the NAc (Supplementary Fig. 5) was used to stimulate VTA GABA projections, which would induce a pause response in CINs during each CS $^+$ presentation (20 Hz of 10-ms pulses for a duration of 300 ms, starting 80 ms after CS $^+$ onset), reproducing the conditioned pause that is known to develop to salient stimuli⁹. During training, GAD-Cre $^+$ and GAD-Cre $^-$ mice developed freezing behaviour, but this did not differ between groups (Fig. 4b). Freezing to the context where foot shocks were delivered was also evident in both groups, but again did not differ between groups (Fig. 4c). Thus, mice were able to learn that the training environment was associated with an aversive event, but forcing a pause of NAc CINs did not affect the associative strength of contextual cues. After training, we placed the mice in a new context and, in the absence of further foot shocks or optogenetic manipulation, asked what had been learned about the CS $^+$ and CS $^-$ presentations by measuring freezing responses to the two stimuli. Because multiple stimulus-foot-shock pairings were trained, GAD-Cre $^-$ mice showed generalized freezing, being unable to discriminate between the CS $^+$ and the CS $^-$ stimuli (Fig. 4d). Notably, GAD-Cre $^+$ mice were able to discriminate effectively between the two stimuli, showing higher levels

of freezing to the CS $^+$ than the CS $^-$ stimuli. Forcing a pause of NAc CINs during the learning of a stimulus–outcome association therefore enhanced detection of the CS $^+$ as the motivationally important stimulus, allowing GAD-Cre $^+$ mice to respond appropriately to the CS $^+$ and avoid generalizing to the CS $^-$ during the subsequent test session.

To examine the possibility that antidromic activation of VTA GABA neurons could inhibit dopamine neurons, thereby influencing our behavioural findings, we first recorded VTA GABA neurons while activating terminals in the NAc. Antidromic action potentials were found in 1 out of 13 GABA neurons of the VTA (Supplementary Fig. 3). Because even a small fraction of interneurons could inhibit dopamine neurons via divergent collaterals, we also monitored VTA dopamine neuron firing while shining light into the NAc. A total of 6 out of 23 dopamine neurons showed responses ranging from inhibition to excitation (Supplementary Fig. 6). Importantly, however, the average dopamine neuron response did not significantly differ from baseline activity. We also performed a behavioural experiment to ensure that blue-light stimulation was not aversive per se, by replacing the foot shock with blue light (Fig. 4e). Contextual and cued freezing did not differ between GAD-Cre $^+$ and GAD-Cre $^-$ mice, and freezing was not controlled by cue presentations (Supplementary Fig. 7). Taken together, these data show that GPN activation in the NAc does not produce an aversive state.

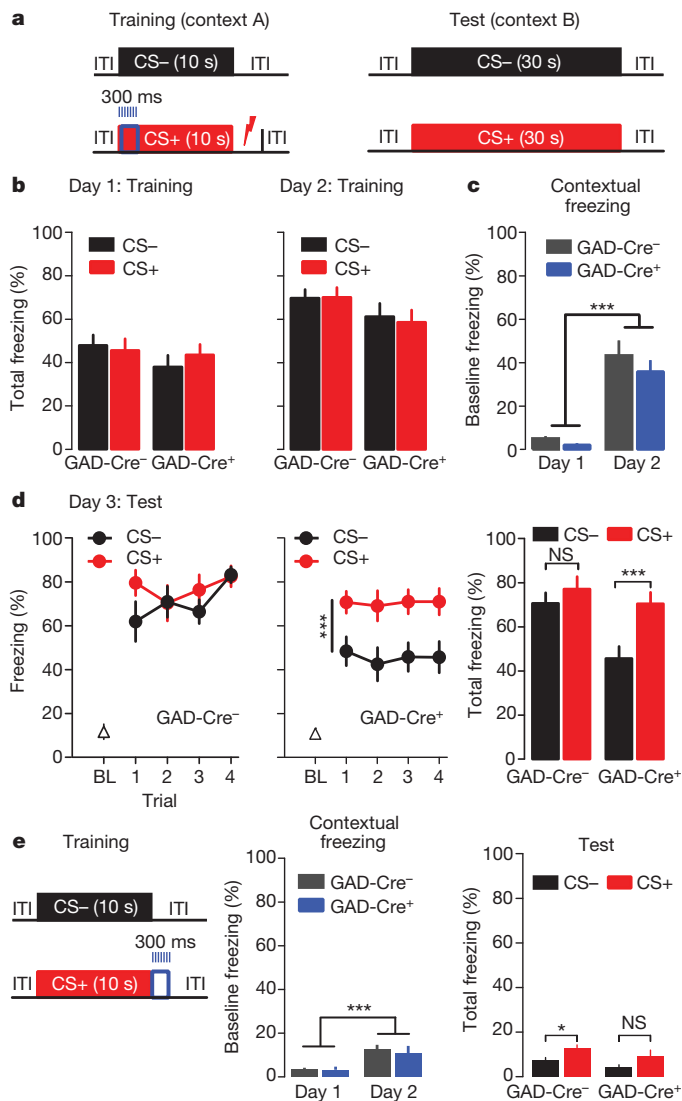


Figure 4 | Activation of VTA GPN axons in the NAc enhances stimulus–outcome learning in an aversive paradigm. **a**, For training, GAD-Cre⁺ ($n = 15$) and GAD-Cre⁻ ($n = 12$) mice were presented with an auditory stimulus (CS+, 10 s) followed by a foot shock (2 s, 0.6 mA) and a second auditory stimulus (CS-, 10 s) associated with no outcome (context A). Each stimulus trial was separated by a variable inter-trial interval (ITI; mean = 100 s). 80 ms after CS+ onset, the blue-light laser was pulsed (300 ms, 20 Hz) to activate VTA GPN axons in the NAc. For the test, stimulus-generated freezing responses were assessed in a new context (context B). **b**, On both training days, GAD-Cre⁻ and GAD-Cre⁺ mice showed equal levels of freezing to the CS- and CS+ presentations. **c**, Contextual freezing was assessed by comparing baseline freezing (%) between the two training days. Contextual freezing was evident (day, $***P < 0.001$), but did not differ between GAD-Cre⁺ and GAD-Cre⁻ mice. **d**, In the test, GAD-Cre⁻ mice froze equally to the CS+ and CS- presentations across trials (left panel). This response pattern differed in GAD-Cre⁺ mice (stimulus \times group, $P < 0.01$; group, $P < 0.05$), which discriminated between the two stimuli across trials (stimulus, $***P < 0.001$; middle panel). A comparison of total freezing (%) to the stimuli confirmed that GAD-Cre⁺ mice, but not GAD-Cre⁻ mice, discriminated between the CS+ and CS- (post-hoc, two-tailed t -test, $***P < 0.001$; right panel). NS, not significant. **e**, In a control experiment, GAD-Cre⁺ ($n = 7$) and GAD-Cre⁻ ($n = 7$) mice received blue-light stimulation in place of a foot shock during training (left panel). Freezing increased between training days (day, $***P < 0.001$), but did not differ between GAD-Cre⁺ and GAD-Cre⁻ mice (note that contextual freezing responses were lower compared with mice that received foot shocks). In the test (right panel), freezing during CS+ and CS- presentations did not differ between GAD-Cre⁻ and GAD-Cre⁺ mice. In GAD-Cre⁻ mice, freezing was increased during CS+ versus CS- presentations (post-hoc, two-tailed t -test, $*P < 0.05$; see also Supplementary Fig. 7). Error bars show s.e.m.

The pathways responsible for CIN pausing in the dorsal striatum and NAc may be different^{15,20,21}. For example, a well-established connection between certain thalamic nuclei and CINs²⁰ can drive a pause after an initial excitation burst¹⁵. This burst–pause depends on dopamine and ACh receptor activation, which was not the case in the NAc (Supplementary Fig. 2) in line with anatomical reports that thalamic inputs to the NAc preferentially target MSNs²². Moreover, the responses of CINs recorded in behaving animals are heterogeneous; often a CIN pause occurs in the absence of any initial excitation, but is nearly always followed by rebound excitation⁹. The direct inhibition that we observe between the VTA and NAc could reflect a common route between the midbrain and striatum to explain such CIN pause responses.

The downstream effects of pausing CIN activity on accumbal processing are not fully known, but may provide a salience window in which temporally coincident dopamine signals can influence accumbal neurons²³. Notably, a rebound excitation was observed immediately after light-induced silencing of CIN activity in both our *in vitro* and *in vivo* preparations, which mirrors that observed in electrophysiological recordings of CINs in the dorsal striatum of non-human primates learning stimulus–outcome associations^{8–12} and after direct inhibition of NAc CINs with halorhopsin¹⁴. The rebound excitation of CINs may directly stimulate dopamine release^{24,25}, and is likely to form a crucial part of the intra-NAc processing of information¹⁴.

The present study examined learning related to an aversive stimulus–outcome association. Studies in monkeys suggest that CINs pause in response to a cue that predicts a salient event, regardless of the valence¹⁰. Moreover, some VTA GABA neurons are activated in similar conditions¹⁷. We now show that the VTA GPNs and NAc CINs are connected, which makes it likely that appetitive learning may also be under the control of the VTA GPN–NAc CIN projection. A wealth of literature has demonstrated the importance of NAc cholinergic tone in regulating appetitive behaviours for both food²⁶ and drug reward^{27,28}. However, as activation of VTA GPNs in the NAc does not alter reward consumption in mice⁵, the function of the pathway may be specialized for learning about stimuli that predict rewarding events.

Disruption of saliency processing could result in failure to isolate and recognize motivationally relevant stimuli efficiently, which may cause responses to develop to non-relevant signals (that is, stimulus generalization). Alternatively, perturbed processing could result in some stimuli becoming overly salient, pulling attention away from otherwise important signals. Both possibilities could feature in neuropsychiatric disorders, for example in generalized anxiety in post-traumatic stress disorder or in drug seeking and relapse in addiction. There is already evidence to show that VTA GABA neurons are affected by drug experience^{29,30}, and it will be of great importance to understand how traumatic events or drug exposure may cause neuroadaptive synaptic plasticity of GABA transmission onto CINs.

We identify a selective projection from VTA GPNs onto NAc CINs. Activation of this projection is sufficient to pause ongoing CIN activity through direct inhibition and thus modify behavioural responses to conditioned stimuli, enhancing contrast between cues with differing motivational relevance. In addition to their established role in controlling dopamine function, our anatomical and functional analysis demonstrates that VTA GABA neurons also control the cholinergic system of the NAc.

METHODS SUMMARY

All experiments were reviewed by the institutional ethics committee and approved by the relevant authorities of the Canton of Geneva. GAD-Cre mice were injected bilaterally into the VTA with AAV-floxed-ChR2(H134R)-eYFP (200–500 nl), and fTom^{PVALBcre} mice injected with a non-floxed construct, using standard surgical procedures. Mice were perfused using 2% paraformaldehyde (containing 0.2% glutaraldehyde for electron microscopy experiments). Immunohistochemical procedures on 50 μ m sections were used to reveal immunoreactivity for ChAT, visualized with CY3 on a confocal microscope. Pre-embedding immunoperoxidase staining with DAB revelation was performed for ChAT and GFP, on ultrathin (60 nm) NAc sections. Images were taken at $\times 3,900$ – $46,000$. For *in vitro*

experiments, acute slices were prepared as previously described⁴ and whole-cell patch clamp recordings performed during varying blue-light stimulation protocols. For *in vivo* electrophysiology experiments, mice were anaesthetized and placed in a stereotactic frame with a fibre optic aimed at the VTA. Single-unit recordings of CINs were performed in the NAc during blue-light VTA stimulation and in some cases the neuron was labelled with 1.5% neurobiotin in the recording electrode (the juxtacellular method) to confirm its identity. Auditory cued and contextual conditioning was performed in AAV-floxed-ChR2(H134R)-eYFP-injected GAD-Cre⁺ and GAD-Cre⁻ mice with fibre-optic implants directed at the NAc using standard surgical procedures. Freezing responses were continuously assessed using automated software.

Received 25 May; accepted 2 October 2012.

Published online 25 November 2012.

- Everitt, B. J. *et al.* Associative processes in addiction and reward. The role of amygdala-ventral striatal subsystems. *Ann. NY Acad. Sci.* **877**, 412–438 (1999).
- Mirenowicz, J. & Schultz, W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature* **379**, 449–451 (1996).
- Robinson, T. E. & Berridge, K. C. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Brain Res. Rev.* **18**, 247–291 (1993).
- Tan, K. R. *et al.* GABA neurons of the VTA drive conditioned place aversion. *Neuron* **73**, 1173–1183 (2012).
- van Zessen, R., Phillips, J. L., Budygin, E. A. & Stuber, G. D. Activation of VTA GABA neurons disrupts reward consumption. *Neuron* **73**, 1184–1194 (2012).
- Van Bockstaele, E. J. & Pickel, V. M. GABA-containing neurons in the ventral tegmental area project to the nucleus accumbens in rat brain. *Brain Res.* **682**, 215–221 (1995).
- Margolis, E. B. *et al.* κ opioids selectively control dopaminergic neurons projecting to the prefrontal cortex. *Proc. Natl Acad. Sci. USA* **103**, 2938–2942 (2006).
- Kimura, M., Rajkowski, J. & Everts, E. Tonicly discharging putamen neurons exhibit set-dependent responses. *Proc. Natl Acad. Sci. USA* **81**, 4998–5001 (1984).
- Aosaki, T. *et al.* Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J. Neurosci.* **14**, 3969–3984 (1994).
- Goldberg, J. A. & Reynolds, J. N. J. Spontaneous firing and evoked pauses in the tonically active cholinergic interneurons of the striatum. *Neuroscience* **198**, 27–43 (2011).
- Apicella, P. Leading tonically active neurons of the striatum from reward detection to context recognition. *Trends Neurosci.* **30**, 299–306 (2007).
- Joshua, M., Adler, A., Mitelman, R., Vaadia, E. & Bergman, H. Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J. Neurosci.* **28**, 11673–11684 (2008).
- Kätzel, D., Zemelman, B. V., Buettner, C., Wölfel, M. & Miesenböck, G. The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nature Neurosci.* **14**, 100–107 (2011).
- English, D. F. *et al.* GABAergic circuits mediate the reinforcement-related signals of striatal cholinergic interneurons. *Nature Neurosci.* **15**, 123–130 (2012).
- Ding, J. B., Guzman, J. N., Peterson, J. D., Goldberg, J. A. & Surmeier, D. J. Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron* **67**, 294–307 (2010).
- Rymar, V. V., Sasseville, R., Luk, K. C. & Sadikot, A. F. Neurogenesis and stereological morphometry of calretinin-immunoreactive GABAergic interneurons of the neostriatum. *J. Comp. Neurol.* **469**, 325–339 (2004).
- Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).
- Inokawa, H., Yamada, H., Matsumoto, N., Muranishi, M. & Kimura, M. Juxtacellular labeling of tonically active neurons and phasically active neurons in the rat striatum. *Neuroscience* **168**, 395–404 (2010).
- Apicella, P., Ravel, S., Deffains, M. & Legallet, E. The role of striatal tonically active neurons in reward prediction error signaling during instrumental task performance. *J. Neurosci.* **31**, 1507–1515 (2011).
- Matsumoto, N., Minamimoto, T., Graybiel, A. M. & Kimura, M. Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J. Neurophysiol.* **85**, 960–976 (2001).
- Schulz, J. M., Oswald, M. J. & Reynolds, J. N. J. Visual-induced excitation leads to firing pauses in striatal cholinergic interneurons. *J. Neurosci.* **31**, 11133–11143 (2011).
- Ligorio, M., Descarries, L. & Warren, R. A. Cholinergic innervation and thalamic input in rat nucleus accumbens. *J. Chem. Neuroanat.* **37**, 33–45 (2009).
- Cragg, S. J. Meaningful silences: how dopamine listens to the ACh pause. *Trends Neurosci.* **29**, 125–131 (2006).
- Threlfell, S. *et al.* Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* **75**, 58–64 (2012).
- Cachope, R. *et al.* Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. *Cell Rep.* **2**, 33–41 (2012).
- Pratt, W. E., Spencer, R. C. & Kelley, A. E. Muscarinic receptor antagonism of the nucleus accumbens core causes avoidance to flavor and spatial cues. *Behav. Neurosci.* **121**, 1215–1223 (2007).
- Hikida, T. *et al.* Increased sensitivity to cocaine by cholinergic cell ablation in nucleus accumbens. *Proc. Natl Acad. Sci. USA* **98**, 13351–13354 (2001).
- Witten, I. B. *et al.* Cholinergic interneurons control local circuit activity and cocaine conditioning. *Science* **330**, 1677–1681 (2010).
- Nugent, F. S., Penick, E. C. & Kauer, J. A. Opioids block long-term potentiation of inhibitory synapses. *Nature* **446**, 1086–1090 (2007).
- Padgett, C. L. *et al.* Methamphetamine-evoked depression of GABA_B receptor signaling in GABA neurons of the VTA. *Neuron* **73**, 978–989 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the members of the Lüscher laboratory for discussion and comments on the manuscript. This work was funded by a grant of the Swiss National Science Foundation to C.L. and the National Center of Competence in Research (NCCR) 'SYNAPSY – The Synaptic Bases of Mental Diseases' of the Swiss National Science Foundation. fTOM^{PVALBcre} mice were provided by A. Holtmaat.

Author Contributions M.T.C.B., K.R.T. and E.C.O'C. performed the *in vitro* electrophysiological recordings. M.T.C.B. and K.R.T. performed the *in vivo* electrophysiological recordings. M.T.C.B. performed the fluorescence immunohistochemical experiments and confocal microscopy. I.N. and D.M. conceived and performed the light and electron microscopic experiments. E.C.O'C. performed the behavioural experiments. C.L., M.T.C.B., K.R.T. and E.C.O'C. designed the study, and C.L. wrote the manuscript with the help of M.T.C.B., K.R.T. and E.C.O'C.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.L. (christian.luscher@unige.ch).

CAREERS

TURNING POINT Materials scientist hopes tenure will facilitate new research **p.459**

FUNDING US National Academies suspends popular policy fellowship **p.459**

NATUREJOBS For the latest career listings and advice www.naturejobs.com

SHUTTERSTOCK/SERGEY NIVENS



COLUMN

Guiding lights

The three Nordic winners of this year's *Nature* mentoring awards respond exceptionally to scientific and personal challenges, says **Philip Campbell**.

As a judge of the 2012 *Nature* Awards for Mentoring in Science, this year featuring the Nordic countries, I noticed that the anecdotes from nominators of one of the short-listed candidates seemed rather cooler in their enthusiasm than others. "Ah," said my fellow judges, who hailed from around the region, "they're from the north."

Nominators from the Nordic nations' more northern reaches may have tended to be more muted in their praise, but trainees and colleagues sent ample compliments for all this year's short-listed candidates — especially the winners. All three have mentored students and postdocs who have subsequently flourished. But

trainees' accomplishments, although important, are just one measure of success. The judges place as much emphasis, if not more, on the human qualities of mentors — their willingness to enable trainees to pursue their own interests, generosity with their time and their ability to serve as a role model as a scientist and lab leader.

This year's judges were drawn from Sweden, Norway, Finland and Denmark, as were the 20 entries. As it happens, all the winners hail from Sweden. The €10,000 (US\$12,900) mid-career award for achievements in mentoring is split between Jens Nielsen, a systems biologist at Chalmers University of Technology in Gothenburg, and Anders Hagfeldt, dean of

chemistry at Uppsala University. The winner of the €10,000 lifetime-achievement award is microbiologist and molecular biologist Staffan Normark, now secretary of the Royal Swedish Academy of Sciences in Stockholm, after posts at Umeå University in Sweden, Washington University in St. Louis and the Karolinska Institute in Stockholm.

WISE COUNSEL

One nominator lauded Nielsen's knack for instantly grasping the challenges of a project and assisting trainees in tackling them. "Jens has the talent to be on top of projects within a few minutes, and thereby is able to contribute ►



CENTRE: JAN-LOF YXELL; RIGHT: MARKUS MARCETIC / ROYAL SWEDISH ACAD. SCI.

Anders Hagfeldt (left), Jens Nielsen (centre) and Staffan Normark help their trainees to fulfil their potential.

► rapidly with input on any ongoing projects,” says Jochen Förster, chief technology officer at the Novo Nordisk Foundation Center for Biosustainability at the Technical University of Denmark in Hørsholm. “This allows him to run large research groups and to continuously attract a high number of national and international PhD students and postdocs.”

Hagfeldt, too, is an astute manager: “His delegation of workload ensures that young researchers develop key skills,” said one nominator. “The most important attribute of Anders Hagfeldt as a mentor is the trust he places in his students and researchers,” said another.

Both have found ways to make time, despite hectic schedules. Nielsen “always found time to hear out colleagues,” noted one nominator. “His extremely large network meant that if a project was stuck he would always know somebody could help.” Another noted that Hagfeldt held weekly coffee meetings to discuss lab issues and weekly lunch meetings for scientific themes.

Some of the simplest attributes were among the most valued. Hagfeldt is a great listener, said trainees. “He is very humble and just lets the other person talk and just with some small questions guides the person in the right direction,” said Tomas Edvinsson, a chemist at Uppsala University. “He encourages cooperation internally and with other groups nationally and internationally in an open spirit.” Hagfeldt says that he tries to “be supportive in the daily harsh work of research and to give confidence that sooner or later the work will pay off”.

Some of Nielsen’s key mentoring moments have come when his protégés faced crises. “The most frequent cause of ‘depressions’ and loss of motivation comes from researchers suddenly becoming overwhelmed by the task and sudden feeling of lack of confidence,” he says. “I assist them with breaking down the tasks into smaller units and encourage them daily to carry out small tasks so they eventually feel they had made progress and faith in their own abilities.”

POTENTIAL FOR GREATNESS

Like most great mentors, Normark demonstrates an ability to tease out the best in his co-workers. Scott Hultgren, a former trainee, recalled how he had joined Normark’s lab in Umeå with potential, but few accomplishments

or publications. “He saw something in me that he was able to foster, nurture, inspire and motivate,” said Hultgren, now a microbiologist at the Washington University School of Medicine in St. Louis, and a member of the US National Academy of Sciences, with more than 200 publications. “There are countless examples of this ‘before-Staffan’ and ‘after-Staffan’ paradigm where he has recruited a young scientist, typically with very few metrics that would be predictive of future achievements, and nucleated their career paths in successful trajectories.”

Normark says that his own mentor was a big influence on his mentoring style. “He never put two postdocs on the same project or even the same pathogen, to avoid internal competition,” he recalls, “and constantly started new innovative projects whenever a postdoc left the lab together with his or her project.”

Another nominator praised Normark’s tendency to encourage trainees to pursue their interests even when that came at a cost to his own research programme. That generosity came along with a healthy dose of investigative freedom. “He gave us the impression that he was more interested in novelty than incremental findings,” said the nominator. “He always emphasized creativity over sheer data production.”

Normark also led by example, offering a steady presence in the face of frustrations and setbacks. “We all wanted to be like him and he inspired us to become better scientists,” said a nominator. “His outlook was always positive and calming when times were rough. I never saw Staffan frustrated by a bad result or slow productivity.” That also meant shielding trainees from the stresses of ancillary duties. “He never shared any frustrations with rejections or university politics with his students,” said another nominator. Principal investigators “don’t realize how much the expression of their frustrations impact the life and career decisions of their students and postdocs”. ■

Philip Campbell is editor-in-chief of *Nature* and the founder of the *Nature Awards for Scientific Mentoring*.

TURNING POINT

Julia Greer

BOB PAZ *Julia Greer is a materials scientist creating next-generation nanomaterials for space exploration at the California Institute of Technology (Caltech) in Pasadena. In August, she added an early-career faculty grant from NASA to her growing list of awards, and she will soon find out if she has secured tenure. Greer hopes that tenure will allow her to spend more time playing the piano — and exploring a new research direction.*

You went back and forth between graduate studies at Stanford University in California and a job at Intel. How did that come about?

I applied late to Stanford, and was accepted, but they could not guarantee funding. So I worked for an adjunct faculty member who had Intel funding to build an X-ray diffractometer. Unfortunately, things ended poorly when he lost funding and criticized my work. I ended up working at Intel for two years.

How did those transitions affect your career?

I came into my own at Intel. I learned to champion my own work and became a self-starter. I stayed there until my husband finished his PhD, then I decided not to let one bad experience limit me, so I returned to Stanford to do my PhD with materials scientist William Nix. It was the most meaningful experience in my life. Using nanopillars, we demonstrated that when you reduce the dimensions of a metallic nanocrystal, it becomes stronger — a phenomenon known as smaller, stronger. That sparked a revolution, with many groups starting to make nanopillars out of gold.

Were you set on getting an academic post?

I had never really considered it, but a professor at Caltech encouraged me to apply for a faculty position. Several people said I had no chance because I would be competing against others who had done postdoc work and had a more mature vision. As a result, I totally relaxed. I thought that I would use the interview process to show the world what we have done. I had other faculty interviews, but ended up taking the position at Caltech.

How did you prepare for the position?

I asked that my tenure clock not start until after I had finished my postdoc at the Palo Alto Research Center in California. Caltech was very flexible. While finishing my postdoc, I wrote a proposal for a National Science Foundation Faculty Early Career Development award. I did a lot of preparation, flying to



Caltech to figure out how to do outreach and how to demonstrate broader impacts. I got the award just as I started at Caltech in 2007.

Do you get overwhelmed?

Yes, mostly because I have two young kids. Everybody has to work out how not to go crazy. In teaching time, I am more hands-off in the lab. I take time to work out, and I practise the piano every night. Playing relaxes me and makes me more creative scientifically. One personal goal is to enter the Van Cliburn International Piano Competition for amateurs.

Do you have a big lab to meet your needs?

I have a group of 15 people and don't want more. Everyone deserves to have a mentor who is there for them. But I did create subgroups. Instead of meeting students individually, I see them in groups of two or three, based on their projects.

How do you manage students who don't meet expectations?

If I see that something is not working out, I ask if the student is satisfied with their progress. If necessary, I encourage them to consider other advisers. I make sure that these conversations are non-confrontational and that somebody else is in the room as a witness.

How might tenure change your research?

Tenure is not something anybody takes for granted, but it would be liberating. I would be able to pursue much riskier ideas. For example, I am very interested in making materials for biomedical devices. Now that I have helped to create a new direction in nanomechanics, it is time for me to get out. It is not healthy to stay in the same area for too long. ■

INTERVIEW BY VIRGINIA GEWIN

CAREER DEVELOPMENT

Mentoring analysed

Mutual respect, clear expectations, personal connections and shared values are key for healthy mentoring, finds a study in *Academic Medicine* (S. E. Straus *et al. Academic Med.* <http://doi.org/jzc;2012>). Interviews with 54 medical-school faculty members at the University of California, San Francisco (UCSF), and the University of Toronto in Canada showed that relationships fail because of poor communication, personality clashes or lack of mentoring experience. The best mentors are trustworthy, listen well, help to set goals and have accessible networks. Co-author Mitchell Feldman, a professor of medicine at UCSF, says that trainees should set agendas for mentoring sessions and update their own development plans.

FUNDING

Policy fellowship pulled

The US National Academies has suspended a fellowship programme popular with early-career scientists seeking science-policy careers. Spokesman Bill Skane says that the National Academies Christine Mirzayan Science & Technology Policy Graduate Fellowship Program is seeking new sources of funding, in part because an endowment from the Carnegie Corporation of New York expires this year. So far, the programme's winter-spring 2013 session has been cancelled. The fellowship, launched in 1997, supports about 50 fellows each year and costs between US\$750,000 and \$1 million a year, including direct expenses and stipends.

ADVOCACY

Postdoc leader sought

Cathee Johnson Phillips, executive director of the US National Postdoctoral Association (NPA) in Washington DC, will step down on 30 April. The NPA's board of directors has launched a nationwide search and hopes to hire a new leader before she leaves. Johnson Phillips joined the NPA — which represents some 2,700 US and Canadian postdocs — in September 2008. She has helped to advocate for stipend increases and benefits for postdocs funded by the US National Institutes of Health; launch a best-practices certification programme for institutions; and facilitate the creation of postdoctoral offices and associations at US universities. She aims to maintain a postdoc-advocacy role but has no specific plans yet.

REVIEW OF THE YEAR: 2062

Things can only get better.

BY JOHN GILBEY

All in all, 2062 hasn't been such a bad year by modern standards. The aftermath of the Silicon Valley debacle is still with us, of course. You will remember the initial panic last year when a protocol breach at Sunnyvale-based biotech start-up Viral Sand led to a chemoviral release — one that has yet to be contained. Although it has dropped out of public awareness, outside California at least, an extensive area west of Highway 101 remains evacuated as the trees, turf and shrubs towards Mountain View and Palo Alto slowly transmute into, well, silicon. The only winner so far has been local coffee chain Dave's Dangerous Drinks — whose espresso, hand-crafted for the billionaire tech trade, is currently the only substance known to halt the crystalline mayhem.

The major — if ironic — loss of life after the poorly controlled blowout at Armstrong in January led to Project Exodus being halted, pending the long overdue inquiry into excess profit-taking by key contractors for the Lunar Commission. There is, as yet, no date set for a resumption of transfer flights to the colony and the members of the extremist Malthus Faction are again petitioning the World Council Office of Population Management to reintroduce the compulsory Life Quality Selection Program — or 'Death Lottery' as it is commonly known. Critics of the programme point to the hugely biased focus found in the original algorithm back in 2055, which discredited the project and led to the forced euthanasia of the software development team.

A low point for some 500 million people occurred when equipment failures at the Nevada-Arizona SPA (Solar Protein Assembly) facility in April put most of the NorthAm Directorate on short rations for the summer. The loss of production at the height of the summer bloom caused major distribution issues, which led to food riots in a dozen prefectures. A spokesperson for the operators cited the withdrawal of government funding from their corporate security as a contributory factor, fuelling speculation that sabotage — presumably by hardline Malthusians — was involved. NorthAm initially refused to speculate on the number of deaths resulting from starvation and cannibalism, but later reassured the public that the figure was probably only in the "low millions".

Perhaps more worrying are the still



unconfirmed stories filtering in from CrystalLattice — the secretive global host body for science and environmental data — that corruption of its universally adopted femtoarray storage media has invalidated much of the material from the second half of the twentieth century. At their conference in June, the surviving members of the Librarians against Climate Change pressure group politely suggested that, in retrospect, the UN mandated carbon-sink policy — which consigned most of the world's physical libraries and human-readable data stores to the bottom of the Marianas Trench in the late '40s — may have been slightly premature.

August saw what was, perhaps, the best news of the year when representatives of Heuristic Tectonics were quoted as saying that "the worst is nearly over". Their analysis of global changes following the surprise impact in 2050 of the asteroid now known as Calamity Jane suggests that less than half of the geological after effects have yet to occur. The East African Rift Sea, now an arm of the Indian Ocean, is slowly stabilizing and some commentators predict that a sustainable fishery may one day be developed on its shores — although why the water is quite so radioactive remains a mystery. The impact on continental, and global, weather patterns is also still under debate: the images

of blizzard conditions around Luxor's temples, and the protracted icing of the Nile, during October

will remain fixed in the minds of climatologists for years to come.

Perhaps not surprisingly, EgoVirtuality (EV) has remained the major preoccupation of most of the human race during the year, with the earnings of the major networks putting them well above all three of the remaining nation states in terms of gross revenues — and I use the term 'gross' advisedly. As food supply and quality of life have become further restricted, the demand for more and wilder programming has been rampant.

The finger of blame in the search for scapegoats for our current problems has pointed increasingly towards the embattled science community again this year — assisted by inflammatory immersive EgoVirtuality series' such as *Rogue EcoScientist*, *Killer Sex Geek* and *Science: Greed Machine* whose fans allegedly burnt down eleven scientific research institutes during the autumn season. These acts of digi-hormone-fuelled violence culminated in the attempted assassination of World Council Speaker Dorothy Ulaka-Chen during her oath-taking ceremony in November — following the revelation on a popular EV fan channel that she holds a PhD in biology. No one has yet been charged with any of these offences.

I am writing this review long-hand, having taken one of the underground craft courses in handwriting recently adopted by the remaining pockets of academia. This act is, however, one of necessity rather than style, as the regional broadcast energy grid in northwest Europe has failed — hopefully only temporarily. The glow from the culprit — the crackling, fizzing aurora overhead — is providing adequate light for the exercise, I note with a wry smile. Writing helps me blot out both the gnawing hunger and the screams of those local EV addicts who are trapped in their non-operative headsets and faced with reality for first time in many years.

Looking out from my frost-gritted window, a range of snow-laden hills is just visible beyond the stunted forest of gaunt fuel-wood trees. With the urban glow absent, this scene has probably changed little in millennia — a thought that provides a modicum of hope. We'll have these problems fixed any day now, won't we. Won't we ...? ■

John Gilbey is a science writer living in west Wales. His nightmares are his own, but he probably needs to talk to someone about them.

JACEY

natureOUTLOOK

PSORIASIS

20/27 December 2012 / Vol 492 / Issue No 7429



Cover art: Andrew Baker

Editorial

Herb Brody,
Michelle Grayson,
Tony Scully,
Nick Haines,
Kevin Singer

Art & Design

Wes Fernandes,
Nicola Hawes,
Gareth Richman

Production

Karl Smart,
Yvonne Strong, Kelly
Hopkins, Leonora
Dawson-Bowling

Sponsorship

Will Piper, Yvette
Smith, Reya Silao

Marketing

Elena Woodstock,
Hannah Phipps

Project Managers

Claudia Deasy,
Christian Manco

Art Director

Kelly Buckheit Krause

Chief Magazine Editor

Tim Appenzeller

Editor-in-Chief

Phil Campbell

After decades of modest advances, psoriasis research has caught fire. In the past ten years, immunosuppressant drugs developed for other autoimmune diseases, such as rheumatoid arthritis, have had a powerful effect in treating severe cases. What's more, this success underscores the crucial role the immune system plays in psoriasis — and points to new ways to treat this inflammatory skin disease. Digging deeper into the underlying causes of psoriasis has unearthed a host of new drug targets. And several promising treatments are nearing clinical approval (see page S58).

Building on this knowledge, genetic studies have provided clues to which parts of the immune system have gone awry in psoriasis — and hinted at where this enigmatic disease might start (S56).

Indeed, psoriasis has become something of a therapeutic test bed for other immune-mediated and inflammatory conditions. Its visible plaques are easy to measure and biopsy to assess treatment efficacy (S50). Psoriasis research has shown the skin to be a vibrant, immunologically active organ (S52). And the skin is home to a variety of microbial communities — disturbances in the skin microbiome may well be what opens the door to psoriasis in the first place (S60).

Although its most obvious impact is on the skin, psoriasis runs deep. Patients are at increased risk of cardiovascular disease, have a higher incidence of metabolic syndrome, and often get psoriatic arthritis. Dermatologist Henning Boehncke outlines the implications of these effects on general healthcare (S55). Psoriasis takes a psychological toll too. In certain individuals, stress can cause the disease to flare up (S62), but holistic treatments using cognitive behavioural therapy seem to help (S64).

We acknowledge the financial support of Celgene in producing this Outlook. As always, *Nature* has full responsibility for all editorial content.

Michelle Grayson

Senior Editor, Supplements

CONTENTS

S50 DERMATOLOGY

Psoriasis uncovered

There's more to it than meets the eye

S52 IMMUNOLOGY

A many layered thing

The skin provides more than barrier protection

S55 PERSPECTIVE

Don't be superficial

Henning Boehncke reveals the hidden harm psoriasis can do

S56 GENETICS

Deep exploration

What really causes psoriasis?

S58 THERAPEUTICS

Silencing psoriasis

Drugs to tackle the most severe cases

S60 MICROBIOME

The surface brigade

Look after the residents

S62 PSYCHODERMATOLOGY

An emotional response

Don't stress

S64 Q&A

Under their skin

Christopher Griffiths explores patients' psychological problems

COLLECTION

S66 Prevalence of metabolic syndrome in patients with psoriasis: a population-based study in the United Kingdom
S. M. Langan et al.

S73 Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity
L. C. Tsoi et al.

S81 Psoriasis: from bed to bench and back
Ken Garber

S85 Local government
Yvonne Bordon

S86 Skin-deep memory
Lucy Bird

S87 Skin infection generates non-migratory memory CD81⁺ T_{RM} cells providing global skin immunity
X. Jiang et al.

Nature Outlooks are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers' expectations. The boundaries of sponsor involvement are clearly delineated in the *Nature Outlook Editorial guidelines* available at http://www.nature.com/advertising/resources/pdf/outlook_guidelines.pdf

CITING THE OUTLOOK

Cite as a supplement to *Nature*, for example, *Nature* Vol XXX, No. XXXX Suppl. Sxx–Sxx (2012). To cite previously published articles from the collection, please use the original citation, which can be found at the start of each article.

VISIT THE OUTLOOK ONLINE

The *Nature Outlook Psoriasis* supplement can be found at http://www.nature.com/nature/outlook/psoriasis_2012/

All featured articles will be freely available for 6 months.

SUBSCRIPTIONS AND CUSTOMER SERVICES

For UK/Europe (excluding Japan): Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants, RG21 6XS, UK. Tel: +44 (0) 1256 329242. Subscriptions and customer services for Americas – including Canada, Latin America and the Caribbean: Nature Publishing Group, 75 Varick St, 9th floor, New York, NY 10013-1917, USA. Tel: +1 866 363 7860 (US/Canada) or +1 212 726 9223 (outside US/Canada). Japan/China/Korea: Nature Publishing Group — Asia-Pacific, Chiyoda Building 5-6th Floor, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo, 162-0843, Japan. Tel: +81 3 3267 8751.

CUSTOMER SERVICES

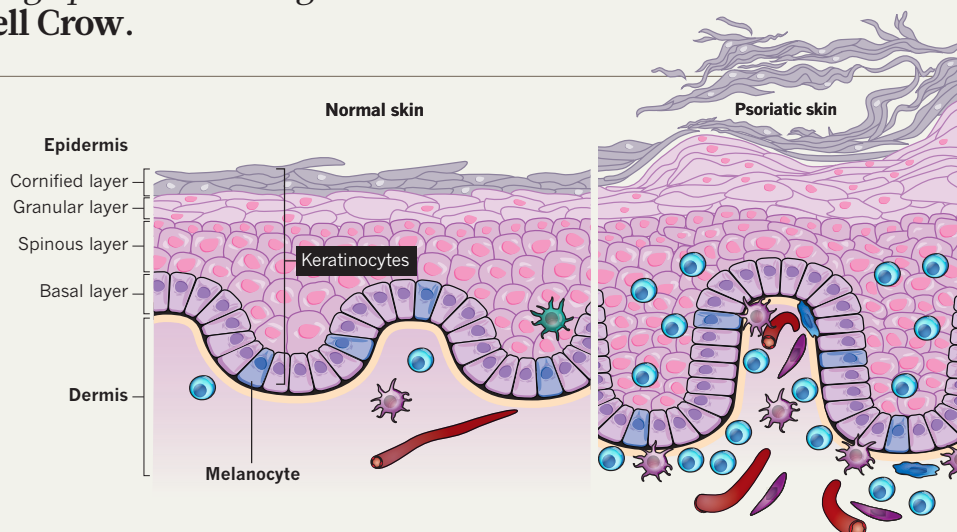
Feedback@nature.com
Copyright © 2012 Nature Publishing Group

PSORIASIS UNCOVERED

Science is finally getting to grips with this enigmatic autoimmune disease. By James Mitchell Crow.

SKIN DEEP

Psoriasis is a non-contagious chronic skin disease affecting the keratinocytes, the cells that predominantly form the epidermis. During outbreaks of the disease, the keratinocytes proliferate at ten times the rate of non-diseased skin cells and fail to mature properly, resulting in raised, inflamed, scaly red skin lesions known as plaques, which can be itchy and painful.

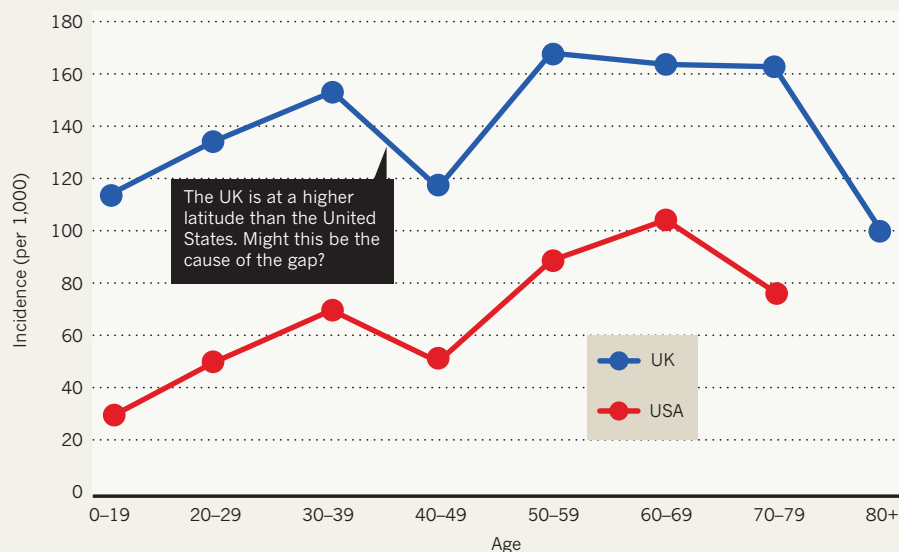


AGE OF ONSET

Although plaque psoriasis is considered a single disease, its severity, progression and response to treatment can vary markedly between patients — suggesting the existence of multiple underlying subtypes. For example, incidence data by age show two spikes, suggesting early- and late-onset forms of the disease.

75%

of patients present with psoriasis before the age of 40



SOURCE: PARISI, R. ET AL. J. INVEST. DERMATOL. DOI:10.1038/JID.2012.339 (2012)

TREATMENT

A broad spectrum of therapies are available to treat psoriasis, depending on the severity of the disease. Newer biologic drugs are assessed by a 'PASI 75' score, representing the percentage of patients achieving at least a 75% reduction in their Psoriasis Area and Severity Index (PASI). Despite considerable progress (see 'Silencing psoriasis', page S58), a cure remains elusive.



>100 YEARS AGO COAL TAR

Tars have been used for nearly 2,000 years to treat skin diseases. They are often messy and smelly, but effective.

1925 GOECKERMAN THERAPY

A combination of coal tar and ultraviolet (UV) irradiation. A course of treatment takes several weeks, and is now less commonly used.

1950s METHOTREXATE

Safe and highly effective for chronic plaque psoriasis. In common use for many years before official FDA approval in 1972.

1951 CORTICOSTEROIDS

Very effective as a short-term treatment. Unsuitable for long-term use because of the growing risk of side effects.

>100 YEARS AGO

1920

1930

1940

1950

1960

PSORIASIS IS COMMON

Global epidemiological data are sparse. There is a lack of uniformity of data collection, so worldwide incidence and prevalence are poorly documented.

Some trends do emerge — for example, prevalence is higher closer to the poles. A combination of genetic and environmental factors are thought to be behind this pattern.

- Temperate (latitude above 40°)
- Subtropical (23–40°)
- Tropical (<23°)

Study type

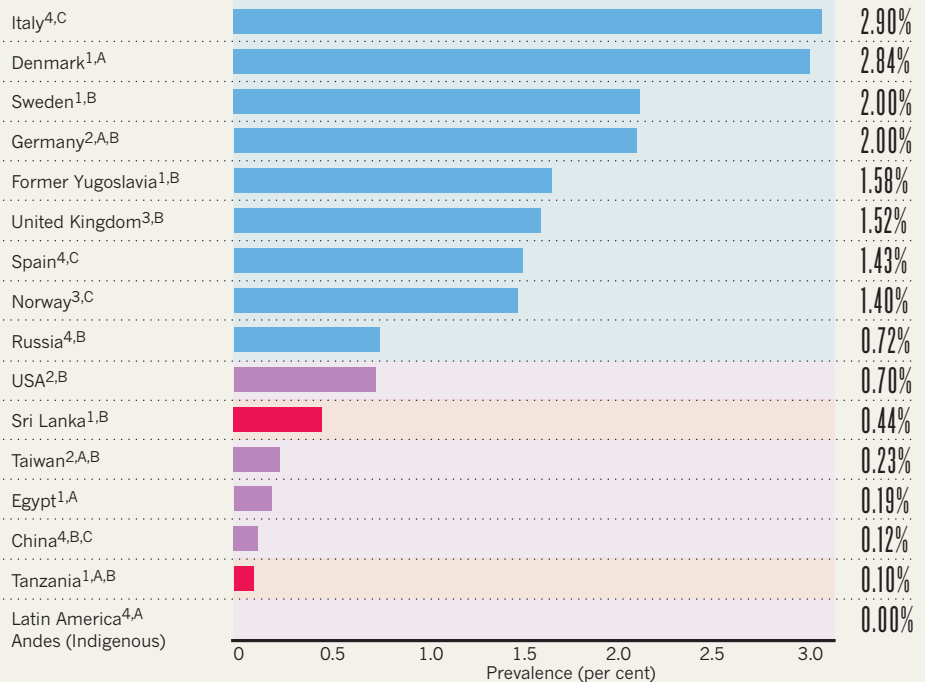
- 1 Point prevalence
- 2 Period prevalence
- 3 Lifetime prevalence
- 4 Not specified

Diagnostic method

- A Dermatologist
- B Physician
- C Self-reported

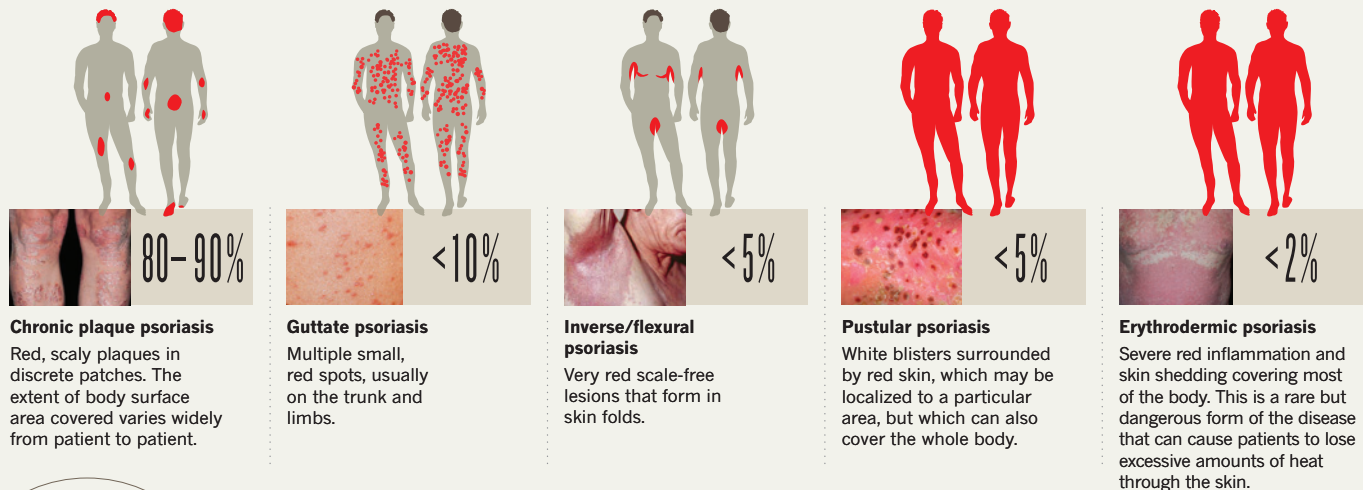
2–3%

average global prevalence of psoriasis



PSORIASIS IS HIGHLY HETEROGENEOUS

Psoriasis comes in several distinct forms. Patients usually display a single form at any one time, although forms can coexist, and one form can be followed by another. Around 80% of psoriasis cases are classified as mild.



200%

increased risk of a heart attack for a 30-year-old with severe psoriasis

1970s

UVB LIGHT

Can clear mild cases of the disease.

PSORALEN AND UVA LIGHT (PUVA)

Possibly slightly more effective, although less convenient, than UVB therapy. A second-line treatment.

1996

ACITRETIN

Effective for generalized pustular and erythrodermic psoriasis, particularly in combination with UVB or PUVA therapy.

1993

VITAMIN D3 ANALOGUES

Well tolerated and effective for long-term use with minimal side effects.

2006

REMICADE (INFLIXIMAB)

PASI 75 at 26 weeks: 50%

2004

ENBREL (ETANERCEPT)

PASI 75 at 24 weeks: 59%

2003

AMEVIVE (ALEFACEPT)

PASI 75 at 14 weeks: 21%

2008

HUMIRA (ADALIMUMAB)

PASI 75 at 16 weeks: 71%

2009

STELARA (USTEKINUMAB)

PASI 75 at 28 weeks: 71–78%

1970

1980

1990

2000

2010



This fruit represents skin, with soft dermis underlying layers of keratinocytes at varying stages of maturity.

IMMUNOLOGY

A many layered thing

No mere passive barrier, the skin is being revealed to be an active part of the immune system. Researchers are now starting to understand its role in driving psoriasis.

BY CLAIRE AINSWORTH

While examining the skin of his psoriasis trial patients, Jim Krueger realized that some of his basic assumptions about the disease were wrong. The year was 1994, and Krueger, then a cell biologist at the Rockefeller University in New York, had been studying psoriasis in order to

better understand skin cancer. The prevailing idea at the time was that excessive division of skin cells caused psoriasis, with chronic inflammation as a side effect. For Krueger, the disease was a logical model for studying skin malignancies, where cell division is key.

As part of the study¹, Krueger's team had given patients a drug that targeted certain immune cells but left skin cells unaffected.

Much to his surprise, their skin markedly improved — in some cases, lesions all but vanished. This effect implicated immune-system cells as the chief instigators of the disease, rather than as secondary players. It also scotched Krueger's idea of using psoriasis as a model for diseases driven by abnormal skin cell division. "I suppose as a scientist if you disprove your hypothesis you've done the best job you can," Krueger observes. "I disproved my hypothesis — and that led me to work on the immunology of skin disease."

The finding was not only a turning point for Krueger's research, but also a major development for psoriasis and skin immunology more broadly. Although evidence had been mounting that the immune system was involved in psoriasis, Krueger's study was the first to discriminate unequivocally between the roles played by keratinocyte cells, which make up most of the skin's physical barrier, and T cells, which help direct the adaptive arm of the immune system. Adaptive immunity can specifically target and remember particular pathogens; in contrast, the innate immune system produces molecules that have a general, non-specific antimicrobial activity.

Over the past two decades, immunologists and psoriasis researchers have shown not only that both arms of the immune system are present in the skin, but that the skin itself is an immunologically active organ. Readily accessible and easy to observe, psoriasis is also shaping up as a model for other chronic inflammatory conditions, such as Crohn's disease and rheumatoid arthritis. Thanks to technological developments in molecular biology and genetics, researchers hope that detailed understanding of the skin's immunology may finally be within reach. "We are going to map chronic inflammation much better than we ever could before," says Frank Nestle, a clinician and immunologist at King's College London.

IMMUNE CELLS

The classical picture of the skin is of a layered barrier that shields the body from the insults of the surrounding world while holding on to moisture and heat. The outer layer, or epidermis, forms the main defence against the outside world; the underlying, cushioning layer, or dermis, contains blood and lymphatic vessels, fat cells and fibroblasts, hair follicles, sweat glands and sensory structures such as touch-sensitive mechanoreceptors (see 'Psoriasis uncovered', page S50).

The epidermis is formed of layers of keratinocytes at varying stages of maturity. Those in the deepest layer are dividing, supplying new keratinocytes that progressively differentiate and rise. By the time they reach the outermost layer they are dead, having become tough, flat, water-retaining husks called corneocytes.

By the 1980s researchers had found that immune-system cells beat a regular path

between the lymphatic system and the skin. As a result, the classical picture of the skin as a passive barrier started to morph into one of an immunologically active organ. Langerhans cells, which are specific to the epidermis but whose function had been unclear for many decades, were recognized as being part of the immune system. In fact, they are a type of dendritic cell that samples the environment and presents fragments of protein to other immune cells as part of the pathogen surveillance system.

The discoveries continued with more kinds of dendritic cell, this time in the dermis. The dermis is now understood to be host to an array of immune cells. Among them are different T-cell types from the adaptive immune system, such as CD4⁺ helper cells that coordinate immune responses and CD8⁺ killer cells that destroy infected, damaged or abnormal body cells. Also present are cells that form part of the innate immune system: natural killer cells, which destroy foreign cells; mast cells, which release inflammation-triggering molecules such as histamine; and macrophages, which gobble up microorganisms and cellular debris. Other dermal immune cells secrete a signalling network of cytokines that coordinate the immune response.

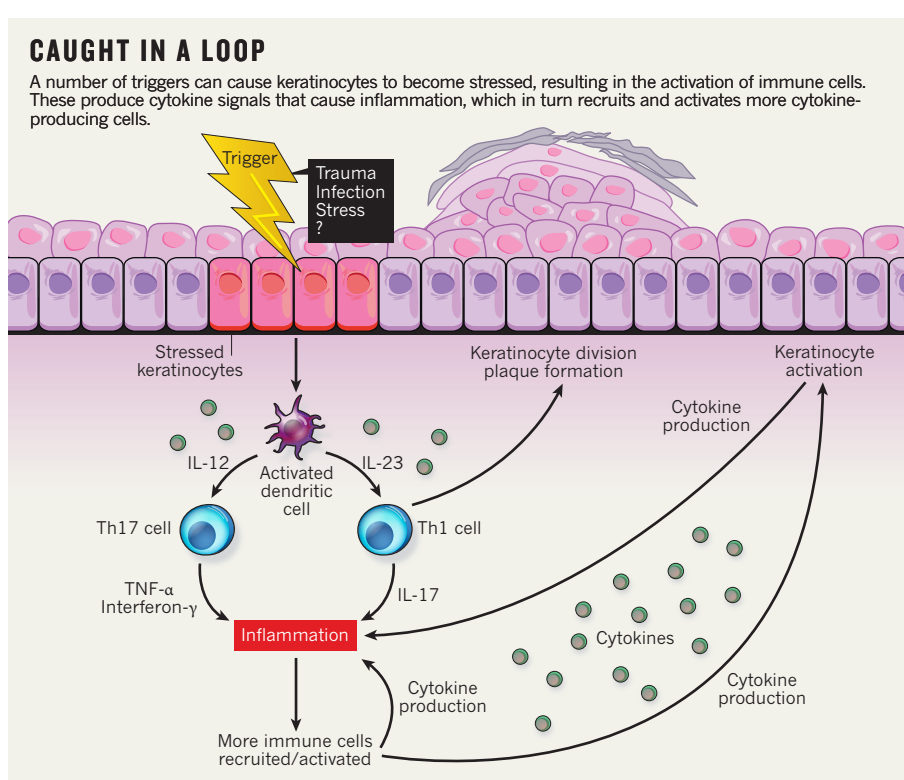
In this milieu, keratinocytes had been thought of simply as the cells that form and renew the skin's barrier. However, in the past few years, it has become clear that they also play a key role in the skin's immune response (see 'Cell summoners').

SYSTEM RESET

If the skin is infected or wounded, the inflammatory response prompts immune cells to produce growth factors that promote both keratinocyte cell division and wound closure. In normal skin, the inflammation resolves once the injury has healed or the infection has cleared. In psoriasis patients, once inflammation has been triggered, it persists. Recent research, plus a hint of serendipity, has helped paint a picture of the major factors involved (see 'Caught in a loop').

Krueger's surprise finding in 1994 revealed the role of T cells in psoriasis. Revelation of the involvement of the tumour necrosis factor (TNF)- α cytokine came from a lucky observation reported² in 2000. A patient receiving treatment for inflammatory bowel disease with a drug that blocks the activity of TNF- α saw her psoriasis dramatically improve. Since then, several TNF- α -blocking drugs, many of which have been co-opted from other autoimmune conditions, have proved successful in treating psoriasis, although they have harsh side effects. The generation of drugs now in clinical trials that target other cytokines (see 'Silencing psoriasis', page S58) is the direct result of this process of studying the immunobiology of psoriasis, says Nestle.

Although these next-generation drugs seem to be good at controlling psoriasis symptoms,



none yet corrects the underlying mechanism: even if treated skin looks healthy, its gene activity can still be abnormal. "You only pull the plug on the inflammatory disease," says Nestle, "you don't reset the immune system."

Resetting the immune system will require a more comprehensive view of how the disease works and what triggers it. With this aim in mind, researchers have turned to 'omics' technologies — genomics, transcriptomics and proteomics — to reveal how biological activity in psoriatic skin compares with that of healthy skin, and how this activity changes after drug treatment.

The transcriptome consists of all the RNA that a cell produces, and is a measure of gene activity. Comparing the transcriptomes of skin cells from psoriatic lesions with those from non-psoriatic skin in the same patient shows differences in gene expression. Several such studies have been published, but they used different patient populations and methodologies, making it hard for researchers to define a core set of common genes simply by looking for overlaps in the genes identified. "There's clearly a complex inflammatory network going on in the skin," says Krueger.

To address this problem, Krueger's team performed a meta-analysis³ on five transcriptomics studies, applying a series of statistical models to the combined data. They identified more than 1,000 genes that were expressed at

different levels in psoriatic skin and normal skin. The team hopes the genes will act as a reference list for future work, as well as providing leads for understanding the pathogenesis of psoriasis and developing new therapies and methods of monitoring response to therapy.

The next big challenge is to make sense of the vast amounts of information being generated. Researchers are turning to computer models to assemble a comprehensive picture of the intricate network of molecular interactions. This systems immunology approach should allow researchers to pinpoint hubs: those molecules that are vital for maintaining a network in a stable configuration. Altering the activity of a hub can abruptly shift the behaviour of a network from one state to another. New drugs that target hubs have the potential to switch a network from a disease-associated state to a healthy one. The result could be treatments that don't just treat the symptoms of psoriasis but actually correct the underlying disease process.

NETWORKING

Although classical genetics and cell biology have provided valuable insights into the interactions in cells and tissues that characterize psoriasis, such approaches are slow and reveal only localized relationships. To build a more complete picture of the whole network, researchers are adopting one of two approaches to sketch out potential molecular interactions and network topologies.

One tack is to identify factors that are regularly found together in the same tissue, indicating a potentially dependent relationship.

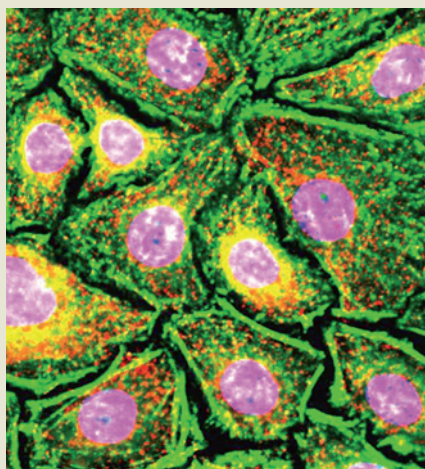
CELL SUMMONERS

Keratinocytes are more than just barrier cells

Since the 1980s, researchers have known that keratinocytes (the cells that make up most of the epidermis) produce antimicrobial peptides — fragments of protein that target and kill disease-causing microorganisms. It is only over the past 5 years, however, that it has become apparent just how big a bag of immunological tricks the keratinocytes have at their disposal.

Keratinocytes express several receptors, including Toll-like receptors (TLRs), which sense the presence of molecules common to many microbes. TLRs play a major role in the innate immune response: the general, non-specific reaction to pathogens. Keratinocytes also release signalling molecules called cytokines that activate other immune cells.

This close confederacy between keratinocytes and the immune system not only defends the skin against infections, but also helps it heal when wounded. In normal skin, cytokines direct the development of keratinocytes to promote skin regrowth and wound closure. In psoriasis, abnormal cytokine secretion retards keratinocyte differentiation, leading



Keratinocytes dyed to show f-actin (green), nuclei (pink) and proteins called defensins (yellow).

to thickened, scaly plaques. “What’s going on in psoriasis is that there is chronic expression of this alternate development pathway,” says Jim Krueger, an immunologist and clinician at the Rockefeller University in New York. “It becomes a self-perpetuating loop.”

However, it is not yet clear how involved keratinocytes are in perpetuating psoriasis. In about 10% of patients, says Krueger, small injuries such as scratches can trigger the formation of new psoriatic plaques, a phenomenon known as the Koebner response. This suggests that keratinocytes have the ability to initiate local patches of chronic inflammation by summoning and activating immune cells.

This ability could be related to the role of keratinocytes in wound healing. In 2012, Krueger’s group reported⁶ that they had mimicked wound healing by separating keratinocytes from other skin cells. By observing the resulting changes in gene expression, they found that keratinocytes not only make cytokines that attract innate immune cells, but they also summon adaptive immune cells, including T cells. “The more we understand about what keratinocytes make,” says Krueger, “the more we can understand about the protective function that the skin has.” **C.A.**

A computer program searches for genes, proteins and other molecules that are coexpressed and that might point to common biochemical pathways. Nestle has used this approach to mine gene-expression data sets to build a psoriasis ‘interactome’ (networks of genes that have similar expression patterns in different tissues and samples). The team hopes the interactome, which they are preparing to publish, will serve as a reference network for interpreting the results of other studies and help uncover new disease mechanisms.

However, this approach will only reveal which molecules are interacting; the next step is to explore the biological consequences of these interactions. This is where the second approach — mathematical modelling — comes in. Nestle’s team, for example, has built a mathematical model⁴ of the immune-cell interactions and cytokine networks in psoriasis. The model shows how changes in cytokine concentrations can alter the inflammatory response. People with a genetic susceptibility to psoriasis secrete more cytokines than those without; Nestle’s model predicts that such a cytokine glut will form a stable network, but one that is prone to suddenly switching states. A trigger, such as a bacterial infection, could tip it into another stable state: chronic inflammation. “In individuals who don’t have the susceptibility, the inflammation resolves,” explains Nestle. “But in patients with psoriasis, it’s stuck in this higher level of equilibrium.”

Nestle’s model also shows how the production of cytokines in patients can deviate from its stable state, triggering oscillations that might lead to flare ups — even in the absence of an environmental trigger. This could explain some of the differences in skin symptoms observed among patients, an idea that biologists can now explore in the lab. These models are “a novel way to create testable hypotheses”, says Nestle.

REVEALING SKIN

One intriguing finding from recent studies of skin immunology is the extent to which psoriasis shares underlying mechanisms with other inflammatory conditions (see ‘Deep exploration’, page S56). Nestle’s team has found⁵ that people with psoriasis have a variant in the gene coding for a receptor for the interleukin-23 cytokine (IL-23). This same variant has also been found in Crohn’s disease, a form of inflammatory bowel disease, and ankylosing spondylitis, a painful, chronic inflammation of the spinal joints. “Psoriasis is a window into other chronic inflammatory diseases,” says Nestle.

Studying psoriasis as a model for chronic inflammation has a number of advantages. One is that the skin is readily accessible. Another is that psoriatic lesions are well defined and accessible: they can be photographed and biopsied, so responses to experimental therapies are easier to quantify than those for internal diseases such as rheumatoid

arthritis. As a result, psoriasis is becoming a proof-of-principle testing ground for new targeted drugs.

In the immediate future, immunologists face the challenge of navigating the deluge of data in order to unravel the remaining mysteries about the role of the immune system in psoriasis. Many of the studies conducted so far focus on established disease. Much less is known about what causes the faulty immune reaction in the first place. What is the molecular trigger for the condition? To what extent are the molecular mechanisms behind psoriasis shared with other conditions? And can the system ever be properly reset?

Krueger believes we are only beginning to reveal the skin’s dynamic role. “There’s a tremendous amount to be learned,” he says, “not only about psoriasis, but also about skin immunology.” There is a long way to go in the skin’s scientific makeover. ■

Claire Ainsworth is a freelance science writer based in Southampton, UK.

1. Gottlieb, S. L. *et al.* *Nature Med.* **1**, 442–447 (1995).
2. Oh, C. J., Das, K. M. & Gottlieb, A. B. *J. Am. Acad. Dermatol.* **42**, 829–830 (2000).
3. Tian, S. *et al.* *PLoS One* **7**, e44274 (2012).
4. Valeyev, N. V. *et al.* *PLoS Comput. Biol.* **6**, e1001024 (2010).
5. Cargill, M. *et al.* *Am. J. Hum. Genet.* **80**, 273–290 (2007).
6. Kennedy-Crispin, M. *et al.* *J. Invest. Dermatol.* **132**, 105–113 (2012).

PERSPECTIVE



Don't be superficial

Severe psoriasis carries cardiovascular risks. Dermatologists should consider more than just patients' outer layers, argues **Henning Boehncke**.

With a prevalence of about 2–3% in the Western world, psoriasis is a common disease¹. Its clinical hallmark is the presence of well-defined red, scaly plaques, which can cover all parts of the body. But the influence of psoriasis extends beyond the skin — the disease's negative impact on psychological and physical well-being is as dramatic as that of cancer, diabetes, arthritis or any other major chronic medical condition². Around a quarter of patients eventually go on to develop psoriatic arthritis, a progressive inflammatory joint disease that causes pain and joint damage and eventually results in disability. And there is mounting evidence connecting psoriasis with adverse cardiovascular and metabolic conditions. Too few doctors — dermatologists and general practitioners alike — are aware of these systemic effects, and how to deal with them.

MOUNTING EVIDENCE

Although epidemiologic studies have long documented an association between psoriasis and various other health problems, definitive evidence of increased morbidity and mortality surfaced only around six years ago³. The effect is most pronounced with cardiovascular disease, and is also connected to severity: patients with severe psoriasis have a higher rate of myocardial infarction than those with mild disease. However, the increased rate of mortality might not be due to cardiovascular disease alone. Rather, psoriasis patients accumulate cardiovascular risk factors such as diabetes, hypertension and obesity, all of which are part of what is called metabolic syndrome. And this syndrome is more common among psoriasis patients than non-psoriatic people.

Evidence that psoriasis is an independent cardiovascular risk factor is mounting. Epidemiologic studies document at least a 50% increase in cardiovascular risk in patients with psoriasis compared with controls⁴. Psoriasis patients have been shown to carry elevated levels of biomarkers for inflammation and cardiovascular risk, such as C-reactive protein; in fact, the more severe the case of psoriasis, the higher the level of this protein⁵. The coronary arteries of psoriasis patients are more often calcified (and to a more profound extent) than the arteries of those without psoriasis⁶. Similarly, the aorta and peripheral blood vessels of psoriasis patients show signs of atherosclerosis and are significantly stiffer than the arteries and veins of non-psoriatic individuals⁷. These associations seem to be restricted to those patients with the most severe psoriasis — patients who receive or have received systemic medication, have undergone phototherapy, or have had in-patient treatment. Patients with mild disease do not generally face a substantially increased risk.

Although the specific mechanisms through which psoriasis impairs cardiovascular health are still under debate, the linkage itself has been firmly established and thoughts must turn to therapies. It might be

possible to reduce atherosclerosis and cardiovascular risk by treating psoriasis with systemic anti-inflammatory drugs. However, that prospect remains distant. What doctors can do now is improve the basic management of patients with severe psoriasis.

CHANGING CLINICAL PRACTICE

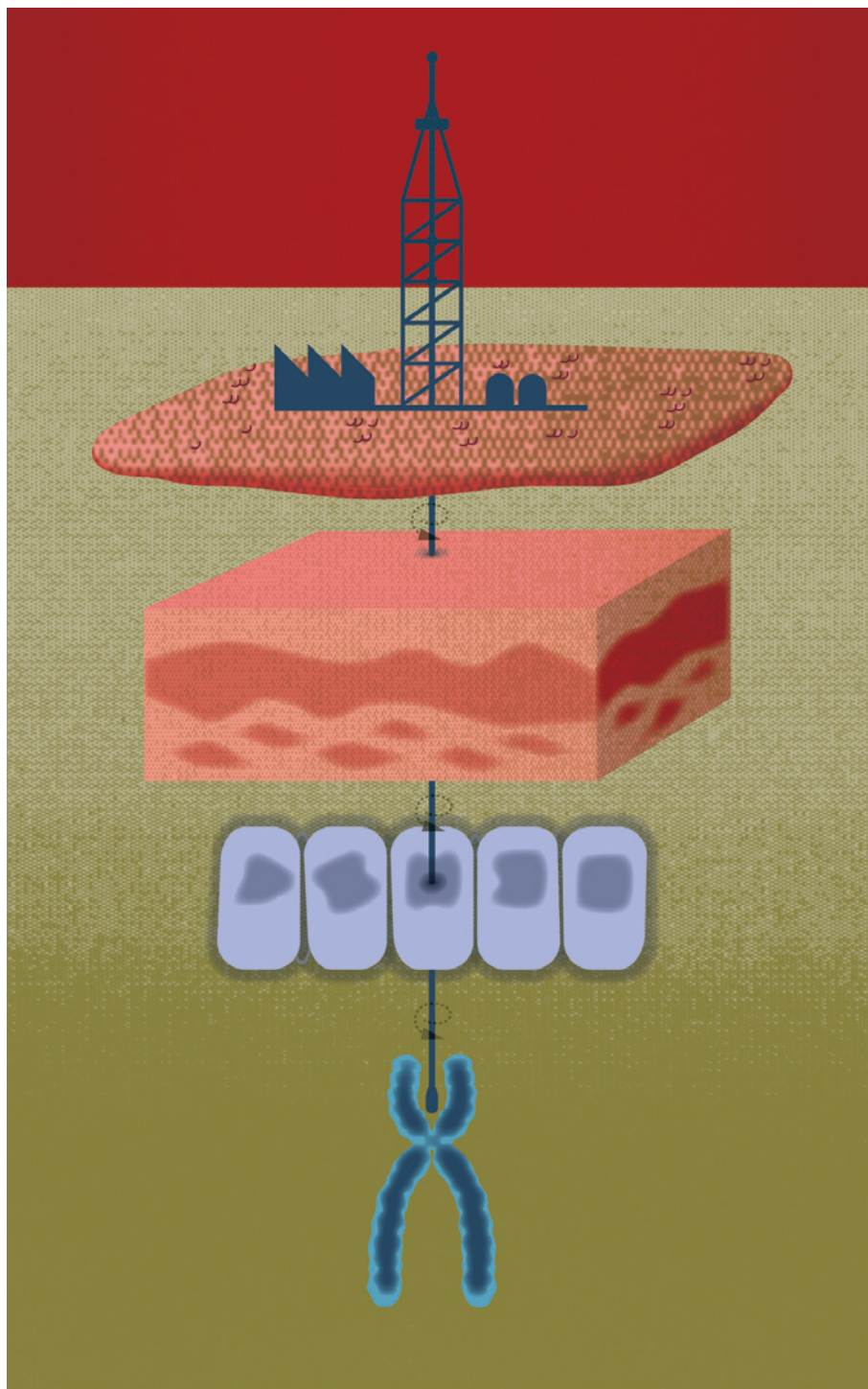
The National Psoriasis Foundation, a research and advocacy organization based in Portland, Oregon, asked its advisory board to compile a checklist to assess patients' cardiovascular risk profile. This checklist was published⁸ in 2008, and comprises a few simple but telling items: pulse rate, blood pressure, blood lipid and glucose levels, and either body mass index or some other gauge of a person's weight. This is a manageable routine for any dermatology practice.

It is not just dermatological specialists who need to be aware of the link between severe psoriasis and cardiovascular risk. Family doctors see patients on a more regular basis, and set targets for blood pressure and lipid levels. These physicians need to realize that psoriasis predisposes a patient to cardiovascular disease. A blood pressure of 140/90 mmHg is tolerable in a patient whose only risk factors are smoking and obesity. However, if this patient also suffers from severe psoriasis, blood pressure should be 120/80 mmHg to yield a comparable cardiovascular risk. Similarly, the therapeutic target for blood lipids is influenced by the number of cardiovascular risk factors. As in the case of blood pressure, lipid levels must be more rigorously controlled if the patient displays more risk factors.

It remains an open question whether there are pathogenetic links between psoriasis and atherosclerosis, or whether long-term systemic therapy can lower cardiovascular risk. Clinicians do not need to wait for certainty on these points. Dermatologists should start assessing their psoriatic patients for cardiovascular risk factors, and doctors more generally should take psoriasis into account when defining therapeutic targets for conditions such as hypertension and hyperlipidaemia. Failure to do so will put lives at risk. ■

Wolf-Henning Boehncke is chairman and head of the Department of Dermatology at Geneva University Hospital, Switzerland.
email: wolf-henning.boehncke@hcuge.ch

1. Nestle, F. O. *et al.* *N. Engl. J. Med.* **361**, 496–509 (2009).
2. Rapp, S. R. *et al.* *J. Am. Acad. Dermatol.* **41**, 401–407 (1999).
3. Gelfand, J. M. *et al.* *J. Am. Med. Assoc.* **296**, 1735–1741 (2006).
4. Gelfand, J. M. *et al.* *Arch. Dermatol.* **143**, 1493–1499 (2007).
5. Coimbra, S. *et al.* *J. Eur. Acad. Dermatol. Venerol.* **24**, 1386–1394 (2010).
6. Ludwig, R. J. *et al.* *Br. J. Dermatol.* **156**, 271–276 (2007).
7. Karadag, A. S. *et al.* *Int. J. Dermatol.* **49**, 642–646 (2010).
8. Kimball, A. B. *et al.* *J. Am. Acad. Dermatol.* **58**, 1031–1042 (2008).



GENETICS

Deep exploration

Recent discoveries are redefining the role of the immune system in psoriasis, and may help to unravel the mystery of the disease's origins.

BY KEN GARBER

Psoriasis genetics took a big leap last month when the number of genome regions known to confer susceptibility to the disease jumped from 21 to 36. The new regions¹ are pieces in the unfinished puzzle that is psoriasis biology, and how researchers assemble these pieces could change our view of the disease.

The new genetic signals strongly suggest a more prominent role in psoriasis for the innate immune system, the body's first line of defence against pathogens. And some of these new psoriasis genes also hint at an answer to a bigger question: where does the disease begin?

AN ARRAY OF RISK

The immune system has two arms, innate and adaptive. The innate immune system reacts quickly and locally to a broad range of invaders, doesn't require prior exposure to them, and lacks memory. Inflammation is an innate response, for example. Adaptive immunity develops over the course of a lifetime. Its long-lived T cells and antibodies are primed from previous exposure to infections, perpetually circulating, ready to strike. Over the past two decades, adaptive immunity had come to be considered the main driver of psoriasis. That conclusion stemmed in part from the abundance of T cells and T-cell-activating dendritic cells found in psoriatic skin, and in part from the effectiveness of drugs that target adaptive immunity, such as ustekinumab (marketed as Stelara by Janssen Biotech of Horsham, Pennsylvania; see 'Silencing psoriasis', page S58). Innate immunity's role in psoriasis was considered secondary.

That's starting to change. Between 2007 and 2011, genome-wide association studies identified 21 psoriasis susceptibility genes, including six involving innate immunity. It was "surprising", says Richard Trembath, a geneticist at King's College London, "how much the genetic evidence pointed towards the involvement of the innate immune system."

To uncover even more susceptibility genes, the Wellcome Trust Case Control Consortium — a group of labs across the United Kingdom — created the Immunochip. This custom genome-sampling device is arrayed with more than 200,000 single nucleotide polymorphisms (SNPs, minuscule genetic changes that vary among individuals) implicated in 12 immune-related diseases. Such diseases were known to have multiple genes in common, and investigators strongly suspected that there were more to be found. More than 10,000 people with psoriasis, along with more than 20,000 control individuals, submitted blood samples for DNA testing.

The Immunochip results hint at an even greater role for innate immunity: at least 5 of the 15 new genomic regions associated with psoriasis risk are involved in innate immunity. (Some regions have not yet been narrowed down to a single gene.) A degree of genetic

ANDREW BAKER

contribution for innate immunity was to be expected, because the skin is a complex, active site of immunity (see 'A many layered thing', page S52). Finding five regions, however, was a surprise, says JT Elder, a geneticist at the University of Michigan, Ann Arbor, and an Immunochip study investigator.

NEXT OF SKIN

Immunochip results also showed that, genetically, psoriasis overlaps most with diseases of the gut. This is not as peculiar as it might seem — skin and the intestines both form a barrier against pathogens and are composed mostly of epithelium. "The skin is epithelium from the outside in, and basically the gut is the same as skin from the inside out," Elder says. "In either case you're dealing with microbes."

Both the skin and the gut marshal similar immunological weapons to repel pathogens. For example, epithelial cells in the intestine and keratinocytes in the skin both release protein fragments called defensins, part of the innate host response to bacterial invaders. Defensins punch holes in bacterial cell walls, eventually killing them, but can also damage host cells if they get out of control. In psoriasis, defensins are produced in vast numbers, causing tissue destruction in the epidermis. As inflammation also stimulates defensin production, this perpetuates a vicious cycle of damage, inflammation and more defensins.

It's not just the Immunochip results that argue for a more important role of innate immunity. In 2011 a French group reported the discovery² of the gene responsible for roughly 10–20% of generalized pustular psoriasis, a severe life-threatening form of the disease. The gene, *IL-36RN*, codes for an anti-inflammatory cytokine that normally turns off innate immune pathways in multiple cell types, including keratinocytes (the cells that constitute almost all the epidermis; see 'Psoriasis uncovered', page S50). When both copies of the gene are mutated, the cytokine can't turn these pathways off, leading to unbridled inflammation that can engulf much of the body. Unlike the psoriasis susceptibility genes identified to date that merely increase risk of the disease, *IL-36RN* mutations cause this extreme form of psoriasis outright. This direct connection indicates that dysregulation of innate immunity is sufficient to cause psoriasis.

Yet more evidence that the innate immune system is responsible for psoriasis comes from work over the past two decades by Anne Bowcock, a geneticist at Washington University in St Louis, Missouri. In 1994 Bowcock found that a region on chromosome 17 was linked to a rare, hereditary form of psoriasis.

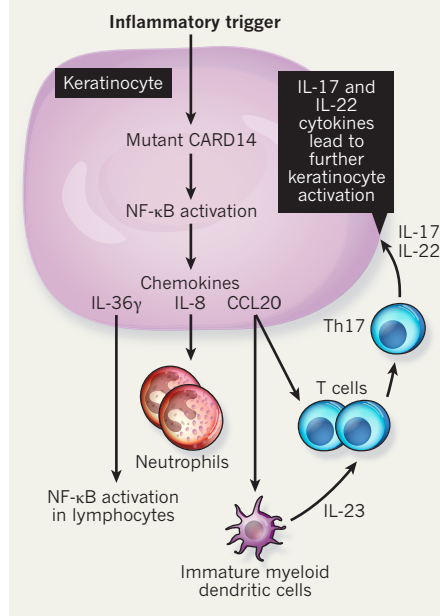
NATURE.COM
For some of the latest research on psoriasis genetics: go.nature.com/uypwky

In terms of biological understanding, however, that knowledge was of limited use: this region, says Bowcock, encompasses more than

70 different genes. To find the specific genetic culprit, Bowcock needed to sequence DNA for the entire region in both affected individuals and controls. In 2012 her work paid off when she reported³ that *CARD14* was the chromosome 17 psoriasis gene. Lab experiments showed that two mutations in *CARD14* cause psoriasis by activating *NF-κB*, a master gene for inflammation. "It's important because we now have two genetic mutations in psoriasis [*IL-36N* and *CARD14*], both involving the innate immune system", that cause psoriasis, says Michelle Lowes, a dermatologist at the Rockefeller University in New York who worked with Bowcock.

INNATE IMMUNITY GONE AWRY

In this model, a skin insult triggers activation of the *NF-κB* pathway via the *CARD14* protein. Signalling proteins cause widespread inflammation, recruitment of T cells and thickening of the skin, all amplified in a feed-forward loop.



Bowcock's team also found a third *CARD14* mutation in a patient with early-onset pustular psoriasis who had no family history of the disease³. This led Bowcock to suspect that, beyond these rare families, *CARD14* might be involved in psoriasis more generally.

Evidence for this broader role emerged from the recent Immunochip psoriasis association study¹ in which *CARD14* was one of the 15 genes identified. "That says that *CARD14* and its pathway are important in the classic form of psoriasis as well," says Bowcock. "What proportion of the classic form isn't clear. But certainly it's part of one of the pathways leading to common psoriasis."

CARD14 is mainly expressed in keratinocytes — a fact that Bowcock says sheds light on the origins of psoriasis. "It's the first instance where we can show that the keratinocyte is likely to play a major role in the development

of the disease, and possibly be where the trigger resides," she says.

STARTING POINTS

Despite the identification of these disease-linked genes, no one knows where psoriasis starts. Some evidence points away from the skin. For example, the first onset of psoriasis often coincides with infection of the throat and tonsils by *Streptococcus* bacteria. The theory is that in priming T and B cells against these invaders, the adaptive immune system confuses some bacterial protein for a host skin protein. (Variants in genes that code for parts of the antigen-presentation machinery have been linked to a predisposition to develop psoriasis, supporting this idea.) This misidentification then leads to an influx of T cells to the skin and hence the autoimmune attack that characterizes psoriasis lesions.

But the many innate immunity genes recently implicated in psoriasis are shifting attention back to the skin. Researchers have found⁴ that most people with psoriasis lack at least one copy of two related genes involved in differentiation of epidermal cells. It is possible that people lacking these genes are unable to properly repair damage to their skin, resulting in a leaky barrier that microorganisms easily pass through. This finding is consistent with the recent Immunochip results. The nature of the innate immunity genes associated with psoriasis, notes Trembath, "tends to suggest the involvement of microbes, particularly viruses".

It's also possible that physical trauma might set off an inflammatory response in keratinocytes that escalates, or that a virus causes an innate immune response and collateral damage to the skin. Both scenarios would involve a cascade of damage and inflammation (see 'Innate immunity gone awry'), accentuated by genetic factors, that draws in the adaptive immune system. "What we're dealing with is a threshold effect for inflammation," says Bowcock. "If the level is too high you will get enhanced immune activation that cannot be turned off."

So attention turns back to the keratinocytes as a potential site for the origin of psoriasis, where a burst of innate immunity later brings in adaptive immunity. "The 'which comes first in psoriasis: abnormalities of keratinocytes or of the immune system' is an ongoing lively debate," says Lowes. "The pendulum has swung in both directions for years." As the genetic profile of psoriasis comes more into focus, the pendulum may be swinging back to the skin. ■

Ken Garber is a freelance science writer based in Ann Arbor, Michigan.

1. Tsoi, L. C. et al. *Nature Genet.* **44**, 1341–1348 (2012).
2. Marrakchi, S. et al. *N. Engl. J. Med.* **365**, 620–628 (2011).
3. Jordan, C. T. et al. *Am. J. Hum. Genet.* **90**, 784–795 (2012).
4. De Cid, R. et al. *Nature Genet.* **41**, 211–215 (2009).



THERAPEUTICS

Silencing psoriasis

The latest drugs hold fantastic promise for people with severe psoriasis. But where are the treatment options for the far larger number with less serious cases?

BY JAMES MITCHELL CROW

There's a buzz among psoriasis researchers. A new generation of therapies are sailing through clinical trials, promising to bring the most debilitating cases of the inflammatory skin disease under control. Not quite a cure, but getting very close.

"It really is a phenomenal period," says Kim Papp, a clinical researcher at Probit Medical Research in Waterloo, Canada. Generating most excitement is a new class of drugs known as the anti-interleukin (IL)-17s. "The anti-IL-17s have demonstrated profound efficacy," says Papp. And he is well placed to judge, having conducted more than 100 clinical trials assessing new psoriasis therapies in the past decade.

Not only are the drugs effective, but so far they seem to be safe. "Given what we've seen in phase II, where all these therapies were well tolerated and continued to be effective, it is very

unlikely that there will be something in phase III to prevent regulatory approval," says Papp.

The field sorely needs fresh treatment options. A survey¹ by the National Psoriasis Foundation, a patient advocacy organization based in Portland, Oregon, found that 40% of people suffering from this disorder are frustrated with the ineffectiveness of their current therapies. Yet most psoriasis patients will not benefit from expensive biologic drugs like the anti-IL-17s. Indeed, none of the antibody-based biologic drugs in development will have an impact on clinical practice for most psoriasis cases in the near term. That distinction is likely to belong to another class of psoriasis drugs entirely, which are proceeding rather more quietly through clinical trials.

COMPETITION TIME

The first biologic therapy for psoriasis was alefacept (Amevive), an anti-T-cell treatment

approved by the US Food and Drug Administration (FDA) in 2003. Five other biologics followed, culminating in ustekinumab (Stelara) in 2009 (see 'Psoriasis through-flow'). These later drugs target different points in the cascade of immune-system signalling molecules, or cytokines, that triggers the skin-cell hyperproliferation characteristic of psoriatic lesions.

Ustekinumab blocks two cytokines, IL-12 and IL-23, which are upstream of IL-17 in this cascade (see 'Deep exploration', page S56). This drug has become the yardstick against which all new biologic psoriasis drugs are measured. Psoriasis therapies are assessed by their PASI 75 score, which represents the percentage of patients achieving at least a 75% reduction in a disease measure called the Psoriasis Area and Severity Index (PASI). At week 12 of treatment, ustekinumab has a PASI 75 of 67% (ref. 2).

As with any drug designed to suppress the immune system, the main concern with ustekinumab is the risk of serious side effects. Such fears are well founded: the development of briakinumab, a biologic with the same mode of action, was halted after a series of major cardiovascular events during clinical trials. Ustekinumab has not been linked with any severe adverse events and, if confidence grows, it might be prescribed more. Severe side effects can take years to become apparent, however. In 2009 a biologic called efalizumab, approved six years earlier, was withdrawn after three patients developed progressive multifocal leukoencephalopathy, a potentially fatal brain disease.

Ustekinumab will soon be in competition with the three anti-IL-17 drugs now in phase III trials: brodalumab, ixekizumab and secukinumab. These agents are proving to be remarkably effective, says Christopher Griffiths, a dermatologist specializing in psoriasis at the University of Manchester, UK. Around half of the phase II trial participants receiving the anti-IL-17 drugs at high dose achieved PASI 100 — that is, complete clearance of psoriasis³⁻⁵. That's a similar proportion to the number who achieve PASI 75 with the current first-line biologics. And as Griffiths points out, "these are patients who have very difficult and recalcitrant disease".

The efficacy results are a surprise, says Rob Kastelein, who works on biologic psoriasis therapies at pharmaceutical giant Merck in Palo Alto, California. Animal studies suggested that blocking the cascade at IL-17 would not be as effective as upstream blockers such as ustekinumab, says Kastelein. "IL-23 is the only molecule where, if you inject it into the skin of a mouse, you get a psoriatic-type lesion." In fact, although ustekinumab targets both IL-12 and IL-23, its efficacy against psoriasis stems entirely from blocking IL-23, Kastelein adds. Armed with this insight, Merck is developing a new agent, MK-3222, that targets only IL-23. It therefore lowers the risk of side effects that might arise from unnecessary interference with IL-12.

Why the anti-IL-17s perform so well remains unclear, says Jonathon Sedgwick, research

ANDREW BAKER

fellow for biotechnology and autoimmunity at Eli Lilly in Indianapolis, Indiana — the pharmaceutical company currently developing ixekizumab. He adds that research into IL-17 is at an early stage. “There are a lot of unknowns.”

Genetic studies hint at a reason. “The anti-IL-17s are impressive for the clinical outcomes, but even more impressive are the changes in gene expression in the psoriatic lesion,” says Andrea Chiricozzi, a research dermatologist at the University of Rome Tor Vergata and the Rockefeller University in New York. Somehow these drugs are downregulating every major cytokine in psoriasis, even those upstream of IL-17 in the cascade. “You have very strong suppression of the psoriasis signature genes — it is even more effective than ustekinumab.”

Potent as they may be, the anti-IL-17 drugs share a limitation with ustekinumab and the other biologics: for the time being, they are aimed only at severe psoriasis.

Psoriasis severity is typically measured by the percentage of a patient's body surface area (BSA) affected by lesions. A BSA of 3–10% is considered moderate, whereas 10% is the traditional cut-off for severe disease. “The patient at 9% is ineligible for these biologic therapies,” says Papp. That's partly down to their cost, and partly down to safety concerns. In the United Kingdom, these treatments can amount to nearly £11,000 (US\$17,500) per patient per year, a cost that insurers and healthcare providers are unwilling to bear for less-than-severe cases. There is also the risk–benefit ratio to consider. As the withdrawn drug efalizumab shows, it can take years to be sure that rare yet severe side effects will not emerge. For milder cases, this risk outweighs the benefit.
















In the meantime, the 90% or so of psoriasis patients considered to have a moderate or mild form of the disease will not have access to the new biologic drugs. “By far the largest proportion of the psoriasis population is bereft of new therapies,” says Papp.

MILD-MANNERED

Two drugs now in phase III trials might bring relief to patients with moderate psoriasis: tofacitinib from New York-based Pfizer (approved by the FDA in November 2012 as Xeljanz — a treatment for rheumatoid arthritis), and apremilast from drug developer Celgene, based in Summit, New Jersey. Both are small-molecule chemical therapies, which should make them cheaper to manufacture than biologic drugs. They can also be taken orally, whereas biologics must be injected. Tofacitinib works by targeting a group of molecules called Janus kinases (JAKs). Knocking out JAKs prevents the body from generating the docking sites needed to receive the cytokines associated with psoriasis. Apremilast, in contrast, works by blocking phosphodiesterase 4 (PDE4), which is essential for cytokine synthesis. “I'm hoping that some of the small molecules, where used as oral agents or introduced as topical agents,

PSORIASIS THROUGH-FLOW

Five biologic drugs are in clinical use. Five more are close behind, along with new oral and topical therapies.

Name (trade name)	Manufacturer	Type	Target	Phase of development
Alefacept (Amevive)	Biogen	Biologic 	T cells (CD2)	Approved 2003
Efalizumab (Raptiva)	Genentech	Biologic 	T cells (CD11a)	Approved 2003, withdrawn 2009
Etanercept (Enbrel)	Amgen	Biologic 	TNF-α	Approved 2004
Infliximab (Remicade)	Janssen Biotech (J&J)	Biologic 	TNF-α	Approved 2006
Adalimumab (Humira)	Abbott	Biologic 	TNF-α	Approved 2008
Ustekinumab (Stelara)	Janssen Biotech (J&J)	Biologic 	IL-12, IL-23	Approved 2009
Brodalumab	Amgen	Biologic 	IL-17 receptor	Phase III
Ixekizumab	Eli Lilly	Biologic 	IL-17	Phase III
Secukinumab	Novartis	Biologic 	IL-17	Phase III
MK-3222	Merck	Biologic 	IL-23	Phase III
Apremilast	Celgene	Small molecule (oral) 	PDE4	Phase III
Tofacitinib	Pfizer	Small molecule (oral) 	JAK	Phase III
Tofacitinib	Pfizer	Small molecule (topical) 	JAK	Phase II
CNTO 1959	Janssen Biotech (J&J)	Biologic 	IL-23	Phase II
AN2728	Anacor	Small molecule (topical) 	PDE4	Phase II

will find their way to treating patients who have less severe disease,” says Papp.

Phase II clinical trial data for both drugs seem underwhelming compared with the results for the biologics. In the tofacitinib study, 67% of patients achieved PASI 75 (ref. 6), while for apremilast the number was 41% (ref. 7). But because no significant side effects were linked to either drug, they could be just what is needed to treat less-severe forms of psoriasis. Apremilast, in particular, might be well suited for treating medium-severity cases, says Papp. One advantage it has is that there are good safety data from PDE4 drugs already in the clinic for other conditions. “Apremilast is potentially a sleeper in the panoply of products.”

Chiricozzi agrees. “With apremilast, you don't have the very high efficacy of the biologic drugs,” he says, “but you do have a very good safety profile.” Although the long-term safety profile of the JAK inhibitors is less clear, they too could ultimately find a market in medium-severity cases, he adds.

Most psoriasis patients, however, have conditions that are too mild to warrant any systemic treatment, whether biologic or small-molecule. The good news for these patients is that it is possible to turn the new small-molecule drugs into ointments. Pfizer, for example, is developing a topical version of tofacitinib and has produced what Chiricozzi says appear to be “very encouraging results”. Although Celgene has not announced the development of a topical apremilast formulation, another company, Anacor Pharmaceuticals, based in Palo Alto, California, does have a topical PDE4 inhibitor in development. Both therapies are

part way through phase II trials.

Creams might seem to be the logical delivery mechanism for psoriasis treatments, but they present unique problems. The skin is a barrier, and it can be difficult to develop an agent that will penetrate the skin in sufficient quantity to have an effect, says Papp. “It takes a lot more to develop a cream than it does to develop a pill.”

Once their safety is established, biologic treatments such as the anti-IL-17s might be prescribed for milder cases. This has happened for biologic drugs in other conditions, says Sedgwick. “But that will take some time.”

Success in developing effective therapies for all psoriasis patients, no matter how mild their condition, would cap what has been a remarkable turnaround in psoriasis drug development. Until ustekinumab was approved, psoriasis medication largely consisted of adopted drugs, from organ transplant immunosuppression to rheumatoid arthritis. And now small-molecule therapies are catching up. “Psoriasis lends itself extremely well to clinical research,” says Griffiths. “Psoriasis is now leading the field, rather than following.” ■

James Mitchell Crow is a freelance science writer based in Melbourne, Australia.

1. Krueger, G. *et al.* *Arch. Dermatol.* **137**, 280–284 (2001).
2. Leonardi, C. L. *et al.* *Lancet* **371**, 1665–1674 (2008).
3. Papp, K. A. *et al.* *N. Engl. J. Med.* **366**, 1181–1189 (2012).
4. Leonardi, C. *et al.* *N. Engl. J. Med.* **366**, 1190–1199 (2012).
5. Rich, P. *et al.* *Br. J. Dermatol.* doi: 10.1111/bjd.12070 (2012).
6. Papp, K. A. *et al.* *Br. J. Dermatol.* **167**, 668–677 (2012).
7. Papp, K. *et al.* *Lancet* **380**, 738–746 (2012).



a natural moisturizer, and others are thought to keep harmful bacteria and viruses from invading. And in the past few years investigators, including Segre, have started to explore whether an imbalance in microbes contributes to skin diseases such as psoriasis and eczema.

SKIN SURVEY

Martin Blaser, a microbiologist at New York University's Langone Medical Center, became interested in psoriasis almost a decade ago after personally developing a mild case. As a physician specializing in infectious diseases, Blaser was well aware of the skin's contingent of normal flora, and wondered whether a change in these microbes had triggered his disease. He investigated this question and in 2007 published one of the earliest surveys of microbes living on the human forearm. Following the launch of the Human Microbiome Project in December 2007, he received a grant from the US National Institutes of Health to further explore the link between microbes and psoriasis along with his then colleague, dermatologist Bruce Strober.

Psoriasis, which typically causes red, raised, scaly, itchy plaques, affects about 2–3% of the world's population (see 'Psoriasis uncovered', page S50). Although Blaser's case was mild, the disease can be debilitating. Those most severely affected "can be among the most miserable patients on Earth", says Strober, now at the University of Connecticut in Farmington. Blaser's team in New York examined the microbes inhabiting the lesions of psoriasis patients. Using a damp cotton applicator, Strober swabbed the plaques for a bacterial sample. Then he took another sample from the patient's healthy skin. He also collected microbes from patients with no signs of the disease.

Blaser discovered³ that the cast of microbes inhabiting the plaque is much more diverse than on healthy skin from either psoriatic or unaffected individuals. He also found that the major phyla of bacteria were present on both psoriatic and healthy skin, but in different ratios. Most notably, there was a striking trend in bacteria from the *Propionibacterium* genus: healthy skin had most, unaffected skin from psoriasis patients had moderate levels, and psoriatic lesions contained least.

This was a small study, says Strober, and it's far from clear whether these differences are a "cause, effect or correlation". Although he admits that "this issue is in its infancy", he is convinced there is a microbial component to psoriasis. One clue, Strober says, comes from identical twins: in 70% of cases, both suffer from the disease. "That's proof that the environment plays a part and it's not all genetic," he explains. What's more, he says, the unexpected flare-ups that characterize the disease might be set off by environmental triggers that alter the composition of the skin's microbe population.

Blaser, meanwhile, has expanded the study and will submit results for publication at the end of 2012. The data suggest there is no single

MICROBIOME

The surface brigade

Our skin is home to thousands of species of bacteria — and when these microscopic societies are disrupted, skin infections can arise.

BY BIJAL TRIVEDI

Each one of us is home to some 100 trillion bacteria. This mass of microbes — accounting for 1–3% of our body weight — resides in our mouth, nose, genitals and intestines, as well as on our skin. Although we are only just beginning to discover what these microscopic hitchhikers do, it is already obvious they are not freeloaders — many play an important role in maintaining human health. There is even preliminary evidence¹ suggesting that when we eradicate certain species of bacteria or alter their relative populations, we inadvertently contribute to the development of diseases ranging from obesity to asthma.

Determining the identity and roles of these microorganisms is the goal of the Human

Microbiome Project. In June 2009, scientists from the project released a survey² of microbes dwelling at 20 locations on the skin of the human body. This census, based on 10 healthy volunteers, produced two significant results, says Julie Segre, a leader of the skin microbiota sequencing efforts at the National Human Genome Research Institute in Bethesda, Maryland. First, it revealed the diversity of microbes that reside on healthy skin. Second, it showed that the organisms varied depending on whether the area was oily like the chest; moist, as in the armpit; or dry, like the forearm (see 'Microbiome map').

Just as the microbes in the gut provide nutritional benefits, those on the skin also earn their keep. Some have evolved to eat dead skin, some transform the oils that skin cells produce into

organism that is present in the psoriatic lesions but not on healthy skin, so there is unlikely to be an obvious microbial villain to exterminate. "It's going to be more complicated," he says.

The link between microbes and psoriasis is still unproven. "It's early days," says Frank Nestle, a skin immunologist and dermatologist at King's College London. Nestle is taking an active role in exploring this association as leader of the UK arm of a US\$6.5-million international consortium to probe the skin microbiome. The project, Microbes in Allergy and Autoimmunity Related to the Skin (MAARS), aims to identify the microorganisms that protect against allergy and autoimmune diseases, as well as the pathogenic bacteria that trigger them. MAARS will focus on the two most prevalent chronic inflammatory skin diseases: eczema and psoriasis.

SURFACE ISSUES

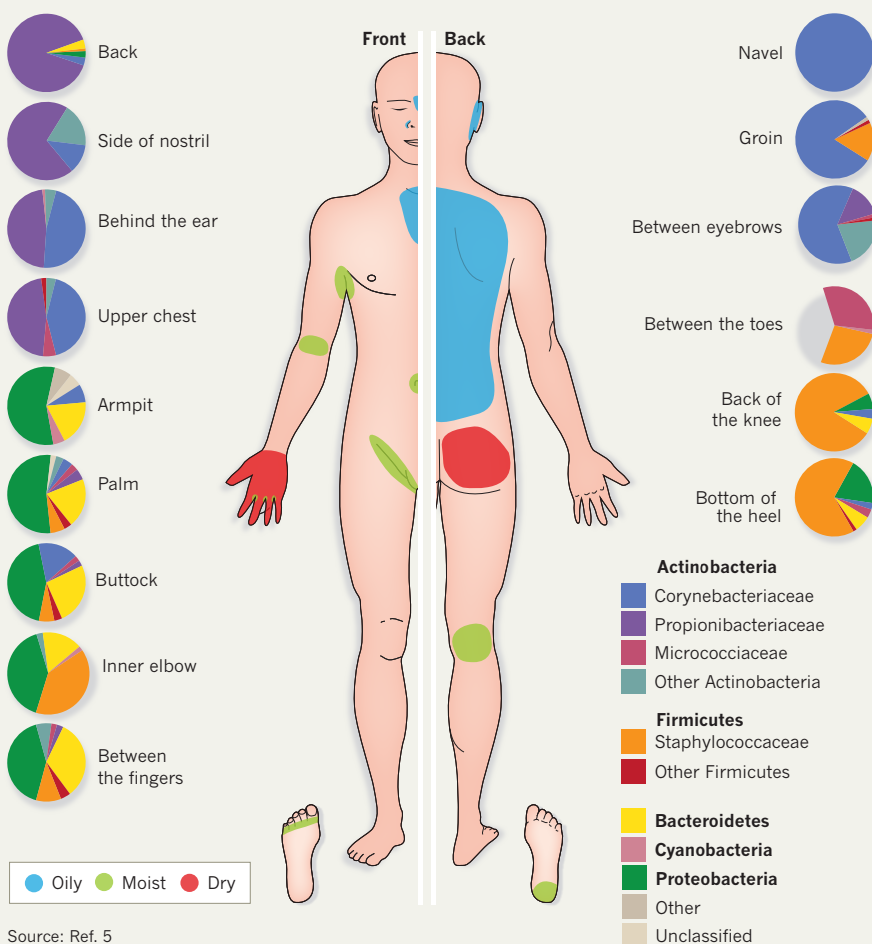
Research on the skin microbiome lags behind that of the gut microbiome. Nevertheless, there are signs that skin microbes play a crucial role in overall health. In 2012, Yasmine Belkaid, an immunologist at the US National Institute of Allergy and Infectious Diseases in Bethesda, described⁴ a link between the microbes on the skin and the body's ability to mount a robust immune response. Belkaid discovered that mice raised in sterile chambers, lacking bacteria on the skin, in the gut and elsewhere, were unable to fight off the parasite *Leishmania major*. However, if the germ-free mice were inoculated with *Staphylococcus epidermidis* — a common skin microbe on mice and humans — at the same time as the *L. major* infection, they were able to defeat the parasite.

Belkaid discovered that *S. epidermidis* interacts with T cells in the skin to produce inflammatory molecules that are central to the immune response. "It's possible that other bacteria on the skin can also rescue this immune response," says Belkaid, "but the point is that there is a link between the community of microorganisms on the skin and the animals' ability to protect against harmful microbes." If a lack of microbes on the skin cripples the immune system, she adds, then it's not unreasonable to think that disrupting the skin's microbial communities might leave it vulnerable to disease.

The evidence linking the microbiota and eczema, which affects more than 15% of children and 2% of adults in the United States, is more solid than for psoriasis. The number of children suffering from eczema has tripled in industrialized nations in the past 30 years, suggesting an environmental trigger. Research has shown⁵ that more than 90% of eczema patients have colonies of *Staphylococcus aureus* on both their affected and unaffected skin. In contrast, the microbe is rarely present on healthy people. Segre adds that, in contrast to psoriatic lesions, eczema sores tend to have a much lower diversity of bacteria than healthy skin. What's more,

MICROBIOME MAP

The human skin is rich with bacteria. The population and ratios vary by region, and depend on the whether the skin site is oily, moist or dry.



Source: Ref. 5

as the condition worsens, *S. aureus* will often overwhelm the whole microbial community.

Proven treatments for eczema include topical antibiotics, steroids and mild bleach baths. But there are no effective antimicrobial treatments for psoriasis. That might seem like a strike against the link between microbes and psoriasis. But, says Blaser, history teaches us not to be so quick to dismiss such an idea. Early efforts to treat gastritis with antibiotics failed, arguing against a bacterial cause. It wasn't until Barry Marshall and Robin Warren discovered *Helicobacter pylori* in the stomachs of patients suffering from inflammation and ulcers that it became possible to develop a targeted regimen — two or three antibiotics plus an acid suppressant — that eradicated the specific bacterium and cured patients.

In a parallel with Blaser's hypothesis about psoriasis, Segre says it is unlikely that *S. aureus* alone leads to eczema. "We have to move away from the idea that one bacterium causes one disease," she says. "It's going to be much more subtle." Blaser and Segre suspect that diseases such as eczema and psoriasis might instead be related to a shift in the balance of microscopic

skin-dwellers. "The challenge is to find out whether there's a causal link between a disturbance in the microbiome and a particular disease," says Blaser.

Future psoriasis therapies, Blaser suggests, might include ointments and creams spiked with particular bacteria that block the growth of other harmful bacteria, or that interact with the immune system to boost its ability to heal skin lesions and maintain skin health.

Whatever eventual treatments come from this work, it is increasingly apparent that our skin is not the first line of defence against the external world: that role falls instead to the army of microbes that live there. Keeping them happy could be the key to keeping our skin soft, supple and healthy. ■

Bijal Trivedi is a freelance science writer based in Washington, DC.

1. Cho, I. & Blaser, M. J. *Nature Rev. Genet.* **12**, 260–270 (2012).
2. Grice, E. A. *et al. Science* **324**, 1190–1192 (2009).
3. Gao, Z. *et al. PLoS One* **3**, e2719 (2008).
4. Naik, S. *et al. Science* **337**, 1115–1119 (2012).
5. Grice, E. A. & Segre, J. A. *Nature Rev. Microbiol.* **9**, 244–253 (2011).



For some people, stress can cause psoriasis symptoms to flare up.

PSYCHODERMATOLOGY

An emotional response

As the link between stress and psoriasis flare-ups becomes clearer, it seems the most vulnerable patients require a new type of treatment.

BY SARAH DEWEERDT

Most people have not-so-fond memories of having their skin break out in pimples during times of teenage stress, such as the night before an exam or their first dance. For many people with psoriasis, stress is written on the skin for their whole lives, and in a much more painful way than a few zits.

Emotional stress is the most frequently cited trigger for psoriasis symptoms, ahead of infections, diet, medication and the weather¹. Not all psoriasis patients have stress-responsive disease, although as many as three-quarters, depending on the particular study population, say that psychological stress causes their disease to flare up.

Dermatologists and psoriasis patients have known about the link with stress for decades. Until recently, however, there was little research to support it. But this is changing. More researchers are probing the connection between stress and psoriasis, and several groups are investigating the cells and molecules involved in the skin's response to stress.

The results are not only illuminating the underlying mechanisms of psoriasis and other inflammatory disorders, but also pointing to the need for a multidisciplinary approach to psoriasis treatment. "Treating a patient with psoriasis is not just treating the skin, it's treating the individual," says Christopher Griffiths, a dermatologist at the University of Manchester, UK, who has conducted multiple studies of the interplay between stress and psoriasis.

In 2009 a group of researchers from Radboud University in Nijmegen, the Netherlands, published² a study that charted day-to-day stresses (such as losing something important, or making a social blunder) and psoriasis symptoms among 62 patients over the course of six months. This was the first large study to follow these relationships as they unfolded over time, as opposed to having patients recall past stresses and psoriasis flare-ups. The study revealed that stressful periods are linked to increased psoriasis symptoms four weeks later. However, this relationship only holds true for patients who worry or scratch a lot; for patients without these characteristics, there was no relationship between stress and the timing of psoriasis symptoms. This has implications for treatment, says clinical psychologist Andrea Evers, a member of the study team. "We have to look at the specific patient profiles that are most vulnerable to stress."

COMMON PROBLEM

Flare-ups of other skin problems, ranging from acne to much rarer disorders such as lichen planus (characterized by an itchy rash on the skin or in the mouth), have also been linked to stress. What do all these conditions have in common?

"Inflammation," says John Koo, director of the Psoriasis, Phototherapy and Skin Treatment Clinic at the University of California, San Francisco. "Inflammation seems to be upregulated by emotional stress." In fact, chronic inflammatory diseases that affect other body tissues, such as rheumatoid arthritis and inflammatory bowel disease, can also be worsened by stress.

One important inhibitor of inflammation is cortisol, a hormone produced by the adrenal gland that is normally released in response to both physical and emotional stress. Research suggests that the regulation of cortisol is dysfunctional in psoriasis, particularly in the large proportion of patients with stress-responsive disease.

One such study³, led by Griffiths, asked psoriasis patients to prepare and deliver a short talk to a panel of strangers — this is the Trier test, a standard technique for inducing social stress. Before the test, some psoriasis subjects reported that stress induces flare-ups, while others said stress had no effect on their disease. After taking the test, patients who reported stress-responsive disease had lower cortisol levels than patients whose disease is not induced by stress.

These findings are not limited to the laboratory. In a six-month study⁴ of stress and psoriasis symptoms, Evers's team also tracked patients' cortisol levels, and found that those with persistently high levels of daily stress had lower average cortisol levels. "These results are in line with the literature that people with, for example, post-traumatic stress disorder or other stress-related disorders also have

GHISLAIN & MARIE DAVID DE LOSSY/GETTY IMAGES

lowered basal cortisol levels,” Evers says.

Although the cortisol response is a good clue to what might be going on, it's not very specific — cortisol levels fluctuate for many reasons. So researchers are investigating how stress affects some cells and molecules thought to play a specific role in plaque formation in psoriasis. Griffiths's team has focused on immune cells called Langerhans cells, which are found near nerve fibres in the upper layers of the skin (see 'Psoriasis uncovered', page S50). “Where the nervous system and the immune system connect is the Langerhans cells,” Griffiths says. “There's direct signalling between the two.”

As psoriatic plaques form, Langerhans cells migrate out of the upper layers of the skin. When the lesions resolve, these cells return. Griffiths has found evidence⁵ that, irrespective of cortisol levels, stress can spur Langerhans cells to move. Among healthy volunteers, the number of Langerhans cells in the upper layers dropped sharply following a Trier test.

Another type of immune cell under investigation is the mast cell. Like Langerhans cells, mast cells are found in large numbers in the upper layers of the skin, near nerve fibres. But whereas the population of Langerhans cells decreases in psoriatic lesions, mast cell numbers increase.

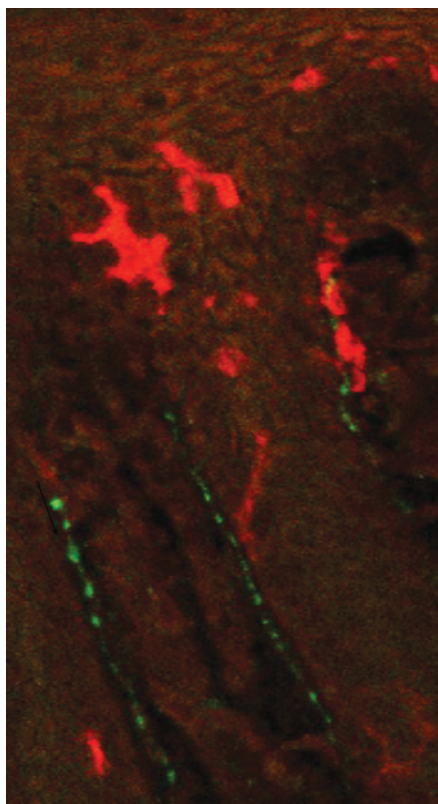
“One of the earliest morphological changes is endothelial swelling and mast-cell activation when a psoriatic lesion develops,” says Ilkka Harvima, a mast-cell biologist and dermatologist at Kuopio University in Finland. “The mast cell participates in the early phases and may direct this entire system.”

Mast cells can be activated by stress hormones such as corticotropin-releasing hormone. Animal studies designed to promote stress — for example, exposing mice to loud noise — show activation of mast cells in the skin. Equivalent studies in humans have not yet been conducted.

Activated mast cells release substances such as histamine that can stimulate the growth of sensory nerve fibres, which are more prevalent in psoriatic lesions. Harvima and other mast-cell biologists think this creates a cycle that helps skin inflammation and psoriasis symptoms persist: stressed nerves release substances that activate mast cells, and these cells encourage the growth of more nerve fibres, which in turn release more of the molecules that activate mast cells.

The effect of stress on psoriasis symptoms is complicated by the fact that psoriasis itself causes stress (see ‘Under their skin’, page S64). People who say their psoriasis is influenced by stress tend to rate their disease as more severe than those who don't, even though the body area covered by psoriasis (the main severity measure doctors use) is the same. Stress responders also tend to have more psoriasis on highly visible areas of the body, such as the face and hands.

There is no evidence that stress causes



There is a close link between nerve fibres (green) and immune cells called Langerhans cells (red).

psoriasis to appear specifically in visible areas, but once these lesions exist “it's really a downward spiral”, Koo says. “The more visible you have it, the more upset people get. The more upset you get, the more inflammatory your skin disease gets.”

CROSS-TALK

The nerve cells, immune cells, skin cells and signalling molecules involved in the cross-talk between the nervous system and the skin are known as the brain-skin axis. The relationship has a deep biological basis, says Mohammad Jafferany, a Michigan-based dermatologist and psychiatrist. “The skin and brain have a common embryological origin,” he points out. Both are derived from ectoderm, the outer layer of cells in an embryo. “This connection is from the very early stages of life.”

Psoriasis is not the only disorder that involves the brain-skin axis, and a better understanding of this disease might point to ways to treat other skin ailments. “The mechanism is probably similar for many inflammatory skin conditions,” says Koo. “Emotional stress exacerbates inflammation relatively non-specifically.” Psoriasis is a particularly good model to use because it is relatively common, has a distinctive appearance, and its severity doesn't change much from day to day (unlike eczema, say).

Psoriasis studies might even yield insights into other chronic inflammatory conditions, such as inflammatory bowel disease. “What we

find with psoriasis and stress is transferable to inflammation in organs that are less accessible,” Griffiths says. “I suspect the mechanisms are very similar.”

The role of stress in exacerbating psoriasis has implications for treatment too. A few studies have suggested that psoriasis patients benefit from interventions such as group therapy, relaxation techniques and stress management.

Cognitive behavioural therapy (CBT), a technique for changing troublesome thoughts and behaviours, has shown promise in relieving stress-responsive psoriasis. A study of 93 patients found that those who received standard psoriasis medications plus a six-week programme of CBT had greater improvements in their psoriasis symptoms than those who received medication alone⁶. What's more, this improvement remained in effect six months after completing the programme. “Once they have those techniques on board, they are more likely to get a prolonged response to therapy,” says Griffiths, who led the work.

CBT “should be implemented more regularly, particularly for highly stress-vulnerable patients”, Evers says. Such patients could be identified through questionnaires and standardized screening instruments soon after psoriasis is diagnosed, so that treatment strategies can then be tailored to their needs.

Such individualized treatment requires new types of collaboration among healthcare professionals. Jafferany, who runs one of a handful of psychodermatology practices in the United States, says that optimal treatment of patients with stress-responsive psoriasis might involve a psychiatrist, a psychotherapist and a social worker as well as a dermatologist.

This multidisciplinary approach is rare. In a survey carried out by a group that included Jafferany, dermatologists in Washington state admitted that although they commonly encounter patients with conditions such as psoriasis that have psychological ramifications, they rarely refer them to a psychiatrist. Furthermore, the survey found that nearly half of the respondents had no formal training in the psychological impacts of psoriasis.

For people with stress-responsive psoriasis, traditional dermatological treatments aren't enough. As Jafferany says: “If you are not looking at the stress element, then you are treating only half the problem.” ■

Sarah DeWeerd is a freelance science writer based in Seattle, Washington.

1. Heller, M. M. *et al. Skin Ther. Lett.* **16**(5), 1–4 (2011).
2. Verhoeven, E. W. M. *et al. Br. J. Dermatol.* **161**, 295–299 (2009).
3. Richards, H. L. *et al. Br. J. Dermatol.* **153**, 1114–1120 (2005).
4. Evers, A. W. M. *et al. Br. J. Dermatol.* **163**, 986–991 (2010).
5. Kleyn, C. E. *et al. J. Invest. Dermatol.* **128**, 1273–1279 (2008).
6. Fortune, D. G. *et al. Br. J. Dermatol.* **146**, 458–465 (2002).



Q&A Christopher Griffiths

Under their skin

Psoriasis can have a profound impact on patients' emotional and social lives. Christopher Griffiths, a dermatologist at the University of Manchester in the United Kingdom, discusses the disease's psychological fallout and its links with stress.

What kind of social and psychological problems do people with psoriasis have?

They have higher levels of depression, anxiety and suicidal ideation than the general population. It's terribly sad. So while people correctly point out that psoriasis is not life-threatening, the disease certainly ruins lives. It was only in the early nineteenth century that leprosy and psoriasis were recognized as separate conditions. Unfortunately, 200 years on, the general public still views psoriasis as contagious. Psoriasis is highly visible, and the immediate response of a lot of people when they see someone with psoriasis is not empathy or sympathy, but to recoil. The average age of onset is in the late teens or early twenties. To have a disease

that carries such a stigma when someone is trying to make their way in life and forge relationships is extremely disabling. And paradoxically, the greater the psychosocial disability the patient suffers, the less likely they are to comply with treatment. It's like a self-fulfilling prophecy. We need to understand why that is the case.

Does psoriasis affect patients' careers?

People believe that psoriasis on the hands or face limits their career choices. They feel that they are judged on their appearance, and therefore may not be put on a front desk as a receptionist, or be allowed to work directly with the public or to serve food. And it's not just a matter of perception. Some people with

severe psoriasis on their hands have difficulty writing and typing, and that might affect their ability to perform in exams or to do their job. The condition can also lead to time off work or school for treatment. All those little increments can add up to a less-than-perfect ability to achieve the heights they would have without the disease.

How do patients deal with the stigma and rejection?

They often practise something called avoidance coping. They will consciously, or subconsciously, avoid situations where others might comment negatively about their psoriasis. For example, a person with psoriasis won't go to a public swimming pool because he or she has heard stories about — or might even

have experienced — being asked to leave by poorly informed members of the public, or even swimming pool staff, who believe that psoriasis is contagious.

There are also profound differences in the way psoriasis patients process other people's responses. We did a study¹ where we took volunteers who had psoriasis and volunteers who didn't and asked them to assign gender to photographs of faces while we scanned their brains using functional magnetic resonance imaging. That task was just a ruse, though. In fact, what we were doing was showing them pictures of disgusted or fearful facial expressions, and ascertaining how they processed them. The interesting thing was that there was no difference between the volunteers with and without psoriasis for expressions of fear, but if they were shown disgust there was a very marked difference — and it was actually the opposite of what we had predicted. In the volunteers with psoriasis, looking at a photo of a disgusted expression was accompanied by reduced blood flow in the anterior insula, the part of the brain that processes disgust. We interpret this as their way of coping, as a shut-off of their response to that disgusted facial expression. And the longer they had had psoriasis, the greater the blood-flow reduction. It's almost a subconscious way of managing the hostile responses of other people.

Do these neurological differences show up in other ways?

We took the same subjects from that first study, sat them in front of a computer screen and showed them pictures of actors simulating different levels of fear or disgust — so, maybe 20% or 30% fear, and so on. We asked the volunteers to press a button when they were certain of the facial expression. With fear, there was no difference between volunteers with and without psoriasis. With disgust, however, a non-psoriatic individual would recognize the expression at about 20% intensity. But even though the volunteers with psoriasis knew what they were looking for, it still took them to about 70% before they were sure they were looking at a disgusted facial expression. So it's a very hardwired coping mechanism.

We see something like this in clinical practice as well. When patients clear their psoriasis outbreaks, their level of tolerance of even small amounts of visible psoriasis in a relapse is far less than it was before. It's as if clearing their psoriasis breaks their coping mechanisms.

How do doctors care for people with a disease that affects so many aspects of their lives?

There has been a revolution in the management of patients with very severe psoriasis, thanks in particular to the development of so-called biologic therapies, such as antibody drugs (see 'Silencing psoriasis', page S58). It's a good example of translational medicine — where understanding the molecular pathways

and immune pathways in a disease has helped us to recognize key targets for these new biologic therapies. The paradox is that you've got fantastic, science-led translational research helping this group of patients with severe disease, but there's a disconnect in understanding how to manage the patient as an individual.

We interviewed² people about what it's like living with psoriasis, and what their concerns are about the management of their condition. In the United Kingdom, there's a lack of engagement between doctors and patients with psoriasis.

“People believe that psoriasis on the hands or face limits their career choices.”

For the most part, doctors look at just the skin and not the disease's effect on the patient as a whole — the links with depression, suicidal ideation and, of course, the emerging evidence of links with cardiovascular disease (see 'Don't be superficial', page S55). The patients do not believe that their doctors understand psoriasis or are willing to give the time to treat the condition properly. I suspect this situation is not unique to the United Kingdom.

The psychosocial problems you've talked about are fairly well established. Aren't doctors aware of the literature?

Doctors know about it, but we're not set up properly to do anything about it. A physician in the UK National Health Service, for example, can only allot about ten minutes to a primary care consultation. That's simply not enough time to determine whether a psoriasis patient is depressed, or to work out what to do about it.

What needs to change?

Doctors need to adopt a holistic approach to psoriasis, to manage it as a complex, long-term condition, in the same way that they now handle higher-profile chronic diseases such as diabetes, rheumatoid arthritis and chronic lung disease. Towards this end, we have launched a five-year study called IMPACT (for Identification and Management of Psoriasis Associated Comorbidity), funded by the UK National Institute for Health Research. The goal of this study is to help healthcare professionals manage psoriasis more effectively — not just treating the skin, but recognizing depression and anxiety better and picking up cardiovascular disease earlier. Our findings will also allow us to develop packages of care — essentially education and counselling — that healthcare professionals can deliver to patients.

What form might psoriasis education and counselling take?

We have developed a cognitive behavioural therapy (CBT) programme for people with psoriasis. It is based on one designed for people coping with pain, disfigurement, social anxiety and stigma. Our CBT programme

educates patients about their condition and gives them cognitive techniques to handle negative beliefs that others might have about them. It teaches coping skills and stress reduction and relaxation techniques, including muscle relaxation.

We have conducted a controlled trial³ to ascertain the utility of the CBT programme. At the end of a six-week intervention, we found that patients who received CBT in addition to their regular topical or systemic therapy had less severe psoriasis than the patients who didn't have the CBT. But, even more importantly, six months after the intervention, that group still had better-controlled psoriasis.

We're also looking at the best way to deliver this kind of holistic care. Does it need to be doctors who do it, or psychologists or counsellors? Or can other healthcare workers deliver this package, and more cost-effectively? One method we've used to deliver this kind of care in a less expensive, less time-consuming way is through an internet-based CBT system funded by the Psoriasis and Psoriatic Arthritis Alliance, where patients with psoriasis can access modules in their own time at home. We have a paper soon to be published in the *British Journal of Dermatology* demonstrating that patients who followed the modules improved their quality of life compared with patients who did not. This improvement was measured using the Dermatology Life Quality Index, which asks a patient to rate how physically uncomfortable his or her psoriasis has been over the past week, as well as how much it has interfered with work, daily activities, relationships and the like.

There may also be an opportunity to use Facebook or other social media, particularly to engage younger people with psoriasis so they can discuss the condition and stay informed of new developments, such as treatments or clinical trials. There are all these different tricks we can use, particularly in the modern era, to manage patients better than we could have done in the past.

How optimistic are you that care for psoriasis patients can be improved?

I'm very optimistic. The reason I say that is because we're starting at a low baseline that offers little psychosocial support for patients. If we can make the management of psoriasis more complete, compelling and holistic, then I think patients are going to benefit. That could improve not just their skin disease, but their general health as well. ■

Interview by Rebecca Kessler, a freelance science journalist based in Providence, Rhode Island.

1. Kleyn, C. E. et al. *J. Invest. Dermatol.* **129**, 2613–2619 (2009).
2. Nelson, P. A. et al. *Br. J. Dermatol.* doi: 10.1111/j.1365-2133.2012.11217.x (2012).
3. Fortune, D. G. et al. *Br. J. Dermatol.* **146**, 458–465 (2002).

Prevalence of Metabolic Syndrome in Patients with Psoriasis: A Population-Based Study in the United Kingdom

Sinéad M. Langan^{1,2}, Nicole M. Seminara², Daniel B. Shin², Andrea B. Troxel^{3,4}, Stephen E. Kimmel⁵, Nehal N. Mehta^{3,5}, David J. Margolis^{2,3} and Joel M. Gelfand^{2,3,4}

Increasing epidemiological evidence suggests independent associations between psoriasis and cardiovascular and metabolic disease. Our objective was to test the hypothesis that directly assessed psoriasis severity relates to the prevalence of metabolic syndrome and its components. A population-based, cross-sectional study was undertaken using computerized medical records from the Health Improvement Network Study population including individuals in the age group of 45–65 years with psoriasis and practice-matched controls. The diagnosis and extent of psoriasis were determined using provider-based questionnaires. Metabolic syndrome was defined using the National Cholesterol Education Program Adult Treatment Panel III criteria. A total of 44,715 individuals were included: 4,065 with psoriasis and 40,650 controls. In all, 2,044 participants had mild psoriasis ($\leq 2\%$ body surface area (BSA)), 1,377 had moderate psoriasis (3–10% BSA), and 475 had severe psoriasis ($> 10\%$ BSA). Psoriasis was associated with metabolic syndrome, adjusted odds ratio (adj. OR 1.41, 95% confidence interval (CI) 1.31–1.51), varying in a “dose-response” manner, from mild (adj. OR 1.22, 95% CI 1.11–1.35) to severe psoriasis (adj. OR 1.98, 95% CI 1.62–2.43). Psoriasis is associated with metabolic syndrome and the association increases with increasing disease severity. Furthermore, associations with obesity, hypertriglyceridemia, and hyperglycemia increase with increasing disease severity independently of other metabolic syndrome components. These findings suggest that screening for metabolic disease should be considered for psoriasis, especially when it is severe.

Journal of Investigative Dermatology (2012) **132**, 556–562; doi:10.1038/jid.2011.365; published online 24 November 2011

INTRODUCTION

The metabolic syndrome is a clustering of cardiovascular risk factors, specifically obesity, hypertension, dyslipidemia, and insulin resistance (Eckel *et al.*, 2005), which has been associated with an increased risk of cardiovascular disease (CVD) beyond traditional risk factors (Mente *et al.*, 2010). The prevalence of metabolic syndrome is increasing in the United States (US; Ford *et al.*, 2002) and partly in Europe,

paralleling the rising prevalence of obesity worldwide (Mokdad *et al.*, 2003; Mente *et al.*, 2010). Systemic inflammation is associated with metabolic syndrome, with T helper type 1 proinflammatory cytokines such as tumor necrosis factor- α and nonspecific measures of inflammation such as C-reactive protein levels being elevated in patients with the syndrome compared with those without (Lakka *et al.*, 2002). However, there is a limited understanding of the relationship between chronic inflammatory diseases and the prevalence of metabolic syndrome.

Psoriasis is the most common T helper type 1 inflammatory disease, affecting more than 125 million people worldwide (National Psoriasis Foundation). The severity of psoriasis in the general population is variable, with most patients having mild disease (Kurd and Gelfand, 2009), defined as involving $\leq 2\%$ of the body surface area (BSA). Epidemiological evidence suggests that psoriasis is associated with an increased frequency of cardiovascular risk factors and adverse cardiovascular outcomes, including myocardial infarction (Gelfand *et al.*, 2006), stroke (Gelfand *et al.*, 2009), and cardiovascular death (Mehta *et al.*, 2010a). Psoriasis, especially if severe, may be a risk factor for atherosclerotic CVD, beyond traditional risk factors (Gelfand *et al.*, 2006, 2009; Mehta *et al.*, 2010a). Moreover, patients with

¹Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK; ²Department of Dermatology, University of Pennsylvania, Philadelphia, Pennsylvania, USA; ³Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pennsylvania, USA; ⁴Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania, USA and ⁵Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence: Sinéad M. Langan, Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK. E-mail: sinead.langan@lshtm.ac.uk

Abbreviations: AIS, Additional Information Services; ATP, Adult Treatment Panel; BMI, body mass index; BSA, body surface area; CI, confidence interval; CVD, cardiovascular disease; GP, general practitioner; NCEP, National Cholesterol Education Program; OR, odds ratio; THIN, The Health Improvement Network

Received 26 April 2011; revised 4 July 2011; accepted 7 August 2011; published online 24 November 2011

severe psoriasis die at an age approximately 5 years younger than patients without psoriasis, with CVD being the most common cause of the excess mortality in these patients (Abuabara *et al.*, 2010). Mechanistic studies of the metabolic syndrome (Shah *et al.*, 2009) and insulin resistance (Mehta *et al.*, 2010b) suggest that chronic T helper type 1 inflammation that characterizes psoriasis, metabolic syndrome, diabetes, and CVD may partly explain the association between these phenotypically distinct diseases.

A number of small, epidemiological studies have reported associations between psoriasis and the metabolic syndrome (Azfar and Gelfand, 2008; Gisondi and Girolomoni, 2009; Al-Mutairi *et al.*, 2010; Mebazaa *et al.*, 2011); however, population-based data in which the severity of psoriasis is objectively determined and individual components of the metabolic syndrome are directly measured are lacking (Augustin *et al.*, 2010). Therefore, our objective was to examine whether there is an association between psoriasis and the metabolic syndrome in a broadly representative population of patients. We also investigate whether the degree of association varies with the extent of skin involvement with psoriasis.

RESULTS

Table 1 describes the demographics of the study population. At the end of the survey collection period, 4,634 of 4,900 provider-based surveys were completed, giving a response rate of 95%. Our cohort included 4,065 people with confirmed psoriasis and 40,650 matched controls. The mean age of psoriasis patients was 1.2 years higher than that of controls ($P < 0.001$), and 51% of psoriasis patients were male compared with 48% of controls ($P < 0.001$). A total of 2,044 (53%) participants had mild psoriasis ($\leq 2\%$ BSA), 1,377 (35%) had moderate psoriasis (3–10% BSA), and 475 (12%) had severe psoriasis ($> 10\%$ BSA). Information on body mass index (BMI), blood pressure, high-density lipoprotein, glucose, and triglyceride levels was available for 41,249 (92%), 44,019 (98%), 25,234 (56%), 28,743 (64%), and 25,067 (56%) of the patients, respectively (Table 2a). These measurements were available in similar numbers of patients with and without psoriasis.

Metabolic syndrome was identified in 34% of participants with psoriasis compared with 26% of controls (odds ratio (OR) 1.50, 95% confidence interval (CI) 1.40–1.61). This association persisted after adjusting for age, gender, and follow-up (adjusted (adj.) OR 1.41, 95% CI 1.31–1.51). Adjusting for smoking and social class did not change the study findings, and these were not retained in the final model. Psoriasis severity affected the degree of association, with the metabolic syndrome seen in 32% of the patients with mild disease (adj. OR 1.22, 95% CI 1.11–1.35), 36% with moderate disease (adj. OR 1.56, 95% CI 1.38–1.76), and 40% with severe psoriasis (adj. OR 1.98, 95% CI 1.62–2.43). Modest but statistically significant interactions were detected between psoriasis and age, and between psoriasis and sex, whereby the OR of metabolic syndrome and psoriasis was slightly higher in the younger age groups and in women (data not shown).

Table 1. Demographic details of the cohort

Characteristic	No psoriasis, <i>n</i> (%)	Psoriasis, <i>n</i> (%)
Overall	40,650 (90.9)	4,065 (9.1)
<i>By psoriasis extent</i>		
$\leq 2\%$	N/A	2,044 (52.5)
3–10%	N/A	1,377 (35.3)
$> 10\%$	N/A	475 (12.2)
<i>Gender</i>		
Men	19,304 (47.5)	2,081 (51.2)
Women	21,346 (52.5)	1,984 (48.8)
<i>Age group (years)</i>		
< 50	13,348 (32.8)	1,082 (26.6)
50–54	10,040 (24.7)	1,000 (24.6)
55–59	9,271 (22.8)	945 (23.3)
> 60	7,991 (19.7)	1,038 (25.5)
Median (interquartile range) of maximum individual measurements		
Body mass index (kg m^{-2})	27.1 (23.7–31.2)	28.4 (24.9–32.9)
Systolic blood pressure (mm Hg)	144 (130–162)	149 (135–168)
Diastolic blood pressure (mm Hg)	90 (80–100)	90 (82–100)
Triglyceride level (mmol/l)	1.7 (1.1–2.5)	1.9 (1.2–2.8)
Cholesterol level (mmol/l)	5.8 (5.1–6.7)	6.0 (5.2–6.8)
High-density lipoprotein (mmol/l)	1.5 (1.2–1.8)	1.4 (1.2–1.7)
Low-density lipoprotein (mmol/l)	3.6 (2.9–4.3)	3.7 (3.0–4.3)
Glucose (mmol l^{-1})	5.4 (4.9–6.4)	5.6 (4.9–6.6)
C reactive protein (mg l^{-1})	5 (3–10)	7 (4–13)
Metabolic syndrome, <i>n</i> (%)	10,515 (25.9)	1,389 (34.2)

On studying the components of the metabolic syndrome, the following factors were found to be more common in psoriasis patients than in controls: obesity in 38% vs. 31% (OR 1.38, 95% CI 1.29–1.48), raised triglyceride levels in 36% vs. 28% (OR 1.49, 95% CI 1.39–1.60), diagnosed hypertension in 31% vs. 28% (OR 1.20, 95% CI 1.11–1.29), and raised glucose levels in 22% vs. 16% (OR 1.44, 95% CI 1.33–1.56; Table 2a).

On studying the fully adjusted model to understand which components of the metabolic syndrome were associated with psoriasis after adjusting for other elements of the metabolic syndrome and age, gender, and duration of follow-up, the strongest association of a component of the metabolic syndrome and psoriasis was found with obesity (adj. OR 1.25, 95% CI 1.16–1.34). The association with

Table 2a. Factors associated with the metabolic syndrome

Factors associated with metabolic syndrome	No psoriasis, <i>n</i> (%)	Psoriasis, <i>n</i> (%)	OR (95% CI)
Body mass index <25 kg m ⁻² ¹	10,744 (28.7)	827 (21.9)	1.0
Body mass index 25–<30 kg m ⁻² ¹	14,143 (37.7)	1,401 (37.1)	1.30 (1.19–1.42)
Body mass index 30–<35 kg m ⁻² ¹	7,678 (20.5)	887 (23.5)	1.52 (1.37–1.68)
Body mass index >35 kg m ⁻² ¹	4,907 (13.1)	662 (17.5)	1.78 (1.59–1.98)
Triglyceride levels ≥1.7 mmol l ⁻¹ ²	11,181 (27.5)	1,453 (35.7)	1.49 (1.39–1.60)
Low HDL (<1.04 mmol l ⁻¹ (men) and <1.29 mmol l ⁻¹ (women)) ³	8,180 (20.1)	1,007 (24.7)	1.32 (1.22–1.43)
Raised BP (systolic ≥130 mm Hg or diastolic ≥85 mm Hg) ⁴	24,187 (59.5)	3,571 (87.9)	1.36 (1.24–1.50)
Hypertension diagnosis	11,204 (27.6)	1,265 (31.1)	1.20 (1.11–12.9)
Type 2 diabetes mellitus	3,445 (8.5)	454 (11.2)	1.36 (1.23–1.51)
High-glucose measurement (>6.1 mmol l ⁻¹) ⁵	6,644 (16.3)	884 (21.8)	1.44 (1.33–1.56)

Abbreviations: BP, blood pressure; CI, confidence interval; HDL, high-density lipoprotein; OR, odds ratio.

¹BMI measured in 41,249: 3,777 (93%) with psoriasis, 37,472 (92%) controls.

²Triglycerides measured in 25,067: 2,545 (63%) with psoriasis, 22,522 (55%) controls.

³HDL measured in 25,234: 2,538 (62%) with psoriasis, 22,696 (56%) controls.

⁴BP measured in 44,019: 4,023 (99%) with psoriasis, 39,996 (98%) controls.

⁵Glucose measured in 28,743: 2,599 (64%) with psoriasis, 26,144 (64%) controls.

Table 2b. Factors including psoriasis extent associated with the metabolic syndrome

Factors associated with the metabolic syndrome	No psoriasis (<i>n</i> , %)	Mild psoriasis (<i>n</i> , %), OR (95% CI)	Moderate psoriasis (<i>n</i> , %), OR (95% CI)	Severe psoriasis (<i>n</i> , %), OR (95% CI)
Body mass index <25 kg m ⁻² ¹	10,744 (28.7)	472 (24.7), 1.0	247 (19.4), 1.0	72 (16.6), 1.0
Body mass index 25–<30 kg m ⁻² ¹	14,143 (37.7)	715 (35.0), 1.15 (1.02–1.30)	483 (35.1), 1.49 (1.27–1.74)	140 (32.2), 1.48 (1.11–1.96)
Body mass index 30–<35 kg m ⁻² ¹	7,678 (20.5)	415 (21.7), 1.23 (1.07–1.41)	316 (24.8), 1.79 (1.51–2.12)	126 (29.0), 2.45 (1.83–3.28)
Body mass index >35 kg m ⁻² ¹	4,907 (13.1)	310 (16.2), 1.44 (1.24–1.67)	228 (17.9), 2.02 (1.68–2.43)	97 (22.3), 2.94 (2.17–4.01)
Triglycerides ≥1.7 mmol l ⁻¹ ²	11,181 (27.5)	686 (33.6), 1.33 (1.21–1.46)	511 (37.1), 1.56 (1.39–1.74)	202 (42.5), 1.95 (1.62–2.34)
Low HDL (<1.04 mmol l ⁻¹ (men) and <1.29 mmol l ⁻¹ (women)) ³	8,180 (20.1)	480 (23.5), 1.22 (1.10–1.35)	348 (25.3), 1.34 (1.19–1.52)	139 (29.3), 1.64 (1.35–2.00)
Raised BP (Systolic ≥130 mm Hg or diastolic ≥85 mm Hg) ⁴	24,187 (59.5)	1,306 (63.9), 1.20 (1.10–1.32)	907 (65.9), 1.31 (1.17–1.47)	329 (69.3), 1.53 (1.26–1.87)
Hypertension diagnosis	11,204 (27.6)	627 (30.7), 1.16 (1.06–1.28)	433 (31.5), 1.21 (1.07–1.35)	151 (31.8), 1.21 (0.98–1.49)
Type 2 diabetes mellitus	3,445 (8.5)	220 (10.8), 1.28 (1.11–1.48)	150 (10.9), 1.30 (1.10–1.56)	58 (12.2), 1.50 (1.14–1.98)
High-glucose measurement (>6.1 mmol l mm Hg) ⁵	6,644 (16.3)	429 (21.0), 1.36 (1.22–1.52)	291 (21.1), 1.37 (1.20–1.57)	129 (27.2), 1.91 (1.56–2.34)

Abbreviations: BP, blood pressure; CI, confidence interval; HDL, high-density lipoprotein; OR, odds ratio.

¹BMI measured in 41,249: 3,777 (93%) with psoriasis, 37,472 (92%) controls.

²Triglycerides measured in 25,067: 2,545 (63%) with psoriasis, 22,522 (55%) controls.

³HDL measured in 25,234: 2,538 (62%) with psoriasis, 22,696 (56%) controls.

⁴BP measured in 44,019: 4,023 (99%) with psoriasis, 39,996 (98%) controls.

⁵Glucose measured in 28,743: 2,599 (64%) with psoriasis, 26,144 (64%) controls.

obesity demonstrated a dose-response relationship, with a 14% increase in obesity in those with mild psoriasis (adj. OR 1.14, 95% CI 1.03–1.27), 34% increase in obesity in those with moderate psoriasis (adj. OR 1.34, 95% CI 1.18–1.53), and a 66% increased odds of being obese in those with severe psoriasis (adj. OR 1.66, 95% CI 1.33–2.07; Tables 2b and 3). Hypertriglyceridemia and hyperglycemia showed similar dose-response relationships with severity of psoriasis.

There was a 20% increased odds of having raised triglyceride levels in individuals with psoriasis overall, independent of obesity (adj. OR 1.20, 95% CI 1.10–1.31). This association also demonstrated an increased odds of having raised triglyceride levels, from 10% in those with mild psoriasis (adj. OR 1.10, 95% CI 0.98–1.25) to 46% in those with severe psoriasis (adj. OR 1.46, 95% CI 1.13–1.88). Raised glucose level was also associated with psoriasis independent of

Table 3. Association between psoriasis severity and components of the metabolic syndrome independent of other components

Psoriasis extent	High blood pressure, OR (95% CI)	Raised triglyceride levels, OR (95% CI)	Low HDL, OR (95% CI)	Hyperglycemia, OR (95% CI)	Obesity (BMI > 30 kg m ⁻²), OR (95% CI)
No psoriasis, n=40,650	1.0	1.0	1.0	1.0	1.0
Psoriasis overall, n=4,065	1.07 (0.96–1.19)	1.20 (1.10–1.31)	0.98 (0.89–1.08)	1.16 (1.06–1.27)	1.25 (1.16–1.34)
<i>By extent</i>					
Mild psoriasis (≤2%), n=2,044	1.03 (0.89–1.20)	1.10 (0.98–1.25)	0.99 (0.87–1.13)	1.11 (0.97–1.26)	1.14 (1.03–1.27)
Moderate psoriasis (3–10%), n=1,377	1.02 (0.85–1.24)	1.31 (1.13–1.51)	0.94 (0.80–1.11)	1.16 (0.99–1.35)	1.34 (1.18–1.53)
Severe psoriasis (>10%), n=475	1.32 (0.91–1.92)	1.46 (1.13–1.88)	1.05 (0.80–1.39)	1.31 (1.00–1.71)	1.66 (1.33–2.07)

Abbreviations: BMI, body mass index; CI, confidence interval; HDL, high-density lipoprotein; OR, odds ratio.

obesity, with a 16% increased odds of raised glucose overall (adj. OR 1.16, 95% CI 1.06–1.27), with the association being strongest in the group with the most severe disease (Table 3).

Sensitivity analysis using the revised Adult Treatment Panel (ATP) III, International Diabetes Federation criteria or by limiting lab values to the first or most recent observation, and excluding individuals on psoriasis treatments that are known to have an impact on components of the metabolic syndrome, e.g., ciclosporin or acitretin, did not significantly change the study conclusions (data not shown).

DISCUSSION

Psoriasis is associated with the metabolic syndrome in a “dose-response” manner, with a 22% increase in the odds of developing the metabolic syndrome in those with mild psoriasis, 56% increase in those with moderate disease, and a 98% increase in those with severe psoriasis. In a fully adjusted model looking at associations between factors comprising the metabolic syndrome and psoriasis after adjusting for other components, independent associations were seen between psoriasis and obesity (25% increased odds), raised triglyceride levels (20% increased odds), and raised serum glucose levels (16% increased odds in a “dose-response” manner from mild to severe psoriasis).

The strengths of this investigation are that it is a large population-based study with a population broadly representative of the UK population in the age group of 45–65 years, which minimizes selection bias and increases the external validity (i.e., generalizability) of the findings. The “dose-response” association detected provides compelling evidence for an association between psoriasis and the metabolic syndrome. Study findings were based on laboratory values and objectively measured disease extent, which allowed observation of findings that are to our knowledge previously unreported. Observational study designs are associated with a number of limitations. These include the cross-sectional nature of this study, which does not allow us to determine which developed first—psoriasis or the metabolic syndrome. Second, we cannot be certain that psoriasis caused the metabolic syndrome; factors including diet, physical inactivity,

alcohol, or genetic predisposition, which have not been evaluated in this study, may be functioning as confounding or effect-modifying factors in this relationship (Davidovici *et al.*, 2010), leading to the possibility of residual confounding. In terms of information bias, two aspects of this study make this an unlikely explanation for the findings: (1) laboratory and clinical values were recorded at similar rates in psoriasis patients and controls as part of routine medical care by general practitioners (GPs) unaware of the hypothesis under study; (2) the persistence of the study findings in the sensitivity analysis restricted to the first laboratory or clinical value per person. Disease severity was determined by asking the GPs to rate the extent of skin involvement with psoriasis into simple discrete categories.

Although previous studies have suggested that UK GPs are reasonably accurate in terms of diagnosing psoriasis (Basarab *et al.*, 1996), direct data on the accuracy of GP assessment of the extent of skin involvement with psoriasis are not, to our knowledge, available. We have previously demonstrated the “construct validity” of this approach in that patients rated by GPs as having higher BSA categories are more likely to require frequent visits for psoriasis and require systemic therapy specific for psoriasis or phototherapy (Seminara *et al.*, 2011). Moreover, we used the same categories used in the epidemiological studies conducted by NHANES and NPF in which patients are asked to rate their degree of skin involvement with psoriasis, suggesting that this approach is acceptable (i.e., “face” validity; Dommasch *et al.*, 2010; Krueger *et al.*, 2001; Seminara *et al.*, 2011). Moreover, these data represent “real-world” data, where the extent of psoriasis has been assessed by hundreds of GPs around the UK and resulted in discrimination of the prevalence of metabolic disorders based on these clinical assessments, demonstrating the usefulness of this approach. Nevertheless, our findings are subject to a form of error (i.e., misclassification of the extent of skin involvement) that would be expected to be non-differential and thus bias toward the null. GPs were asked to assess the BSA of involvement that the patient typically demonstrates; this measure may not be stable over time, although a previous large cohort study demonstrated

that despite various therapeutic interventions the severity of psoriasis for individuals did not generally change over time (Nijsten *et al.*, 2007).

This study significantly advances the existing literature on psoriasis and the metabolic syndrome, as this is the first population-based study to use objective measures of psoriasis severity, direct measurement of the components of metabolic syndrome, and standard criteria for diagnosis of metabolic syndrome. Of special interest is the clear “dose-response” relationship between psoriasis severity and the metabolic syndrome. No previous study has, to our knowledge, shown a directional increase in the association with raised triglyceride levels and increasing psoriasis severity independent of the effects of obesity. The consistency with other study findings (Gisondi and Girolomoni, 2009; Al-Mutairi *et al.*, 2010; Augustin *et al.*, 2010; Love *et al.*, 2011; Mebazaa *et al.*, 2011), presence of a “dose-response” relationship, strong associations, and biological plausibility support some causality, but further mechanistic and longitudinal studies are required (Rothman and Greenland, 2005).

A possible biological mechanism that may account for this association is that the proinflammatory state associated with psoriasis functions as a central driving force for development of the metabolic syndrome. In psoriasis patients, Th1 inflammatory cytokines, e.g., tumor necrosis factor- α , IL-1, and IL-6, are increased in skin and blood (Azfar and Gelfand, 2008). These inflammatory mediators may have a range of effects on insulin signaling, lipid metabolism, and adipogenesis. In addition, inflammation-induced insulin resistance may lead to the development of a systemic insulin-resistant state (Mehta *et al.*, 2010b). Further mechanistic studies will be needed to test this hypothesis.

Study findings demonstrate a strong association between psoriasis and the metabolic syndrome, with increasing psoriasis severity being associated with increasing odds of metabolic syndrome. Increased odds of raised triglyceride levels and serum glucose were seen in individuals with psoriasis independent of the effects of obesity. The results of this study firmly establish that the metabolic syndrome is an important comorbidity with psoriasis, and that vigilance and enhanced screening may be important in psoriasis patients, particularly those with severe disease. Examining the components of metabolic syndrome associated with psoriasis, weight reduction is clearly a key step to prevent CVD; however, our findings also show the importance of screening for the other components of metabolic syndrome, particularly hypertriglyceridemia and raised glucose levels, as these tests are more likely to be abnormal in patients with psoriasis independent of traditional risk factors (such as obesity). Small increases in the individual components of metabolic syndrome have led to an 8% absolute increase in the prevalence of metabolic syndrome overall and a 14% increase in those with severe psoriasis. Further prospective studies are required to determine the directionality of the association between psoriasis and metabolic syndrome and to study other unexplored confounders, including diet, physical activity, alcohol, and genetic factors, which may be important residual confounders in this relationship.

MATERIALS AND METHODS

Study design

We conducted a cross-sectional study using The Health Improvement Network (THIN).

Study population

THIN is a computerized longitudinal general-practice database with demographic data similar to the general United Kingdom (UK) population. THIN has anonymized medical record data on 3.4 million “active” patients followed up for a cumulative 50 million person years, and is broadly representative of the UK population. The THIN database contains demographic details, diagnoses, laboratory results, and prescriptions recorded by GPs, the gatekeepers for medical care in the UK. The version of THIN we used contained data from 413 general practices that use the “In Practice Vision” software. A number of studies have confirmed that THIN data are highly accurate, thus making it ideal for use in epidemiological research (Lewis *et al.*, 2007; Seminara *et al.*, 2011). The cohort was identified from individuals in the age group of 45–64 years with at least one psoriasis Read code (using a previously validated coding algorithm (Seminara *et al.*, 2011)) in the 2 years before the survey. Patients were required to be registered with a general practice contributing actively to Additional Information Services (AIS). AIS practices have an agreement to respond to questionnaires; 55% ($n = 228$) of THIN practices were AIS active at the time of sampling. A total of 4,900 eligible patients with psoriasis diagnostic codes were randomly sampled, and questionnaires were sent to their GPs through AIS to verify the presence of psoriasis and the extent of disease. Up to 10 controls in the age group of 45–64 years were randomly matched to each psoriasis patient based on practice; similar to cases, controls needed to be alive and actively registered with at least one GP visit within 2 years at the time of sampling.

Outcomes

Patients were defined as having psoriasis if their diagnosis was confirmed by a questionnaire completed by their GP. The questionnaire also determined the severity of psoriasis, namely mild psoriasis ($<2\%$ BSA), moderate psoriasis (3–10% BSA), and severe psoriasis ($>10\%$ BSA). This approach has been previously well accepted (Feldman, 2004). Cardiovascular risk factors, specifically BMI calculated using standard formulation (overweight was defined as $\text{BMI} \geq 25 \text{ kg m}^{-2}$ and $<30 \text{ kg m}^{-2}$, obese was defined as $\geq 30 \text{ kg m}^{-2}$), hypertension, hyperlipidemia, smoking, and diabetes mellitus, were identified by the presence of diagnostic Read codes and additional recording and laboratory values in the Additional Health Details portion of the database.

Subjects were defined as having metabolic syndrome using the National Cholesterol Education Program (NCEP) ATP III diagnostic criteria (Expert Panel on Detection, 2001). Using NCEP criteria, a person with metabolic syndrome fulfills three or more of the following criteria: central obesity (determined by a $\text{BMI} \geq 30 \text{ kg m}^{-2}$ in THIN), hypertriglyceridemia $\geq 1.7 \text{ mmol l}^{-1}$, low high-density lipoprotein cholesterol (in men $<1.03 \text{ mmol l}^{-1}$ and in women $<1.29 \text{ mmol l}^{-1}$), high blood pressure ($\geq 130/85 \text{ mm Hg}$) and high fasting glucose level ($\geq 6.1 \text{ mmol l}^{-1}$). Time-varying variables were dealt with by selecting the maximum laboratory value or clinical measurement and using the most recent value for BMI. Conditions

were measured from the patients' start date (defined as the latest of the Vision software or computerization in the practice and registration dates of the patient), whereas the end of the study was defined as the earliest date of transfer out, death, or end of the study period in February 2009.

Study size

We calculated that a sample size of 4,900 would yield 4,190 patients, which would be sufficient to detect increased relative risks of 1.14 for a BMI of $\geq 25 \text{ kg m}^{-2}$, 1.37 for hypertension, 1.71 for hyperlipidemia, and 2.0 for diabetes mellitus, with 80% power, respectively, assuming a two-sided test and a significance level of 0.05, and we were satisfied that such differences would be clinically meaningful.

Statistical methods

ORs and 95% CIs for the association between psoriasis overall and by psoriasis extent were calculated using conditional logistic regression. Multiplicative interaction terms were fitted to assess the effect modification by age and sex. Adjusted ORs were determined by adjusting for confounders including age, sex, and duration of follow-up time in THIN. Other possible confounders that were explored included smoking and social class, which were measured using Townsend scores (Phillimore *et al.*, 1994). Further analyses were undertaken of the association between psoriasis and disease extent and the components of the metabolic syndrome to ensure that the findings were not explained by individual components such as obesity. Sensitivity analyses were undertaken using the revised NCEP ATP III definition (glucose cut point $>5.6 \text{ mmol l}^{-1}$) and the International Diabetes Federation definitions of metabolic syndrome (Zimmet *et al.*, 2005). Sensitivity analyses were also carried out using only the first and most recent laboratory value for each individual and in patients who did not receive psoriasis treatments that may affect blood pressure and lipid levels (i.e., ciclosporin or acitretin). All analyses were carried out in Stata SE10 (Stata Corporation, College Station, TX).

Ethics

This study was approved by the University of Pennsylvania institutional review board and the Cambridgeshire Research Ethics Committee, and was funded by the National Heart Lung and Blood Institute of the NIH.

Role of the funding source

The sponsors had no role in the conduct or interpretation of the study. The corresponding and senior author had full access to all data in the study and had final responsibility for the decision to submit for publication.

CONFLICT OF INTEREST

JMG has received grants from Amgen, Pfizer, Novartis, and Abbott, and is a consultant for Amgen, Celgene, Pfizer, Novartis, and Centocor; DJM is on separate data safety monitoring boards for Abbott and Astellas, which might have an interest in the submitted work in the previous 3 years; the remaining authors state no conflict of interest.

ACKNOWLEDGMENTS

This study was supported by an R01 grant and a Graduate Research Supplement (NMS) grant from the National Heart, Lung, and Blood Institute of

the NIH RO1HL089744 (JMG). The funders had no role in the design, analysis, or interpretation of this research. Dr Langan is funded by fellowships from the British Association of Dermatologists and the National Psoriasis Foundation. Ms Seminara is funded by a Graduate Research Supplement grant from the National Heart, Lung, and Blood Institute of the NIH RO1HL089744. Dr Mehta is funded by a grant from the National Psoriasis Foundation and from the National Heart, Lung, and Blood Institute of the NIH K23 HL097151.

Author contributions

JMG, DJM, NMS, ABT, and SEK were involved in the conception of the research question, planning of the study, and application for funding. DBS extracted the data from the THIN database and assisted with data management and guidance on the use of THIN. Further data management was carried out by SML and NMS. SML, NMS, ABT, and JMG were involved in analysis of the data. SML and JMG drafted the manuscript, which was reviewed by all authors. SML is the guarantor. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

- Abuabara K, Azfar RS, Shin DB *et al.* (2010) Cause-specific mortality in patients with severe psoriasis: a population-based cohort study in the U.K. *Br J Dermatol* 163:586–92
- Al-Mutairi N, Al-Farag S, Al-Mutairi A *et al.* (2010) Comorbidities associated with psoriasis: an experience from the Middle East. *J Dermatol* 37:146–55
- Augustin M, Reich K, Glaeske G *et al.* (2010) Co-morbidity and age-related prevalence of psoriasis: analysis of health insurance data in Germany. *Acta Dermatol Venereol* 90:147–51
- Azfar R, Gelfand J (2008) Psoriasis and metabolic disease: epidemiology and pathophysiology. *Curr Opin Rheumatol* 20:416–22
- Basarab T, Munn SE, Jones RR (1996) Diagnostic accuracy and appropriateness of general practitioner referrals to a dermatology out-patient clinic. *Br J Dermatol* 135:70–3
- Davidovici BB, Sattar N, Prinz JC *et al.* (2010) Psoriasis and systemic inflammatory diseases: potential mechanistic links between skin disease and co-morbid conditions. *J Invest Dermatol* 130:1785–96
- Dommasch E, Shin D, Troxel A *et al.* (2010) Reliability, validity and responsiveness to change of the Patient Report of Extent of Psoriasis Involvement (PREPI) for measuring body surface area affected by psoriasis. *Br J Dermatol* 162:835–42
- Eckel RH, Grundy SM, Zimmet PZ (2005) The metabolic syndrome. *Lancet* 365:1415–28
- Expert Panel on Detection Ea, and Treatment of High Blood Cholesterol in Adults (2001) Executive summary of the third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 285:2486–97
- Feldman SR (2004) A quantitative definition of severe psoriasis for use in clinical trials. *J Dermatolog Treat* 15:27–9
- Ford ES, Giles WH, Dietz WH (2002) Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *JAMA* 287:356–9
- Gelfand J, Dommasch E, Shin D *et al.* (2009) The risk of stroke in patients with psoriasis. *J Invest Dermatol* 129:2411–8
- Gelfand J, Neimann A, Shin D *et al.* (2006) Risk of myocardial infarction in patients with psoriasis. *JAMA* 296:1735–41
- Gisoni P, Girolomoni G (2009) Psoriasis and atherothrombotic diseases: disease-specific and non-disease-specific risk factors. *Semin Thromb Hemost* 35:313–24
- Krueger G, Koo J, Lebwohl M *et al.* (2001) The impact of psoriasis on quality of life: results of a 1998 National Psoriasis Foundation patient-member-ship survey. *Arch Dermatol* 137:280–4

- Kurd SK, Gelfand JM (2009) The prevalence of previously diagnosed and undiagnosed psoriasis in US adults: results from NHANES 2003–2004. *J Am Acad Dermatol* 60:218–24
- Lakka HM, Laaksonen DE, Lakka TA *et al.* (2002) The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA* 288:2709–16
- Lewis J, Schinnar R, Bilker W *et al.* (2007) Validation studies of The Health Improvement Network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf* 16:393–401
- Love TJ, Qureshi AA, Karlson EW *et al.* (2011) Prevalence of the metabolic syndrome in psoriasis: results from the national health and nutrition examination survey, 2003–2006. *Arch Dermatol* 147:419–24
- Mebazaa A, El Asmi M, Zidi W *et al.* (2011) Metabolic syndrome in Tunisian psoriatic patients: prevalence and determinants. *J Eur Acad Dermatol Venereol* 25:705–9
- Mehta N, Azfar R, Shin D *et al.* (2010a) Patients with severe psoriasis are at increased risk of cardiovascular mortality: cohort study using the General Practice Research Database. *Eur Heart J* 31:1000–6
- Mehta NN, McGillicuddy FC, Anderson PD *et al.* (2010b) Experimental endotoxemia induces adipose inflammation and insulin resistance in humans. *Diabetes* 59:172–81
- Mente A, Yusuf S, Islam S *et al.* (2010) Metabolic syndrome and risk of acute myocardial infarction: a case-control study of 26,903 subjects from 52 countries. *J Am Coll Cardiol* 55:2390–8
- Mokdad AH, Ford ES, Bowman BA *et al.* (2003) Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA* 289:76–9
- National Psoriasis Foundation. <http://www.psoriasis.org/netcommunity/learn_statistics> Accessed 22 November 2010
- Nijsten T, Looman CW, Stern RS (2007) Clinical severity of psoriasis in last 20 years of PUVA study. *Arch Dermatol* 143:1113–21
- Phillimore P, Beattie A, Townsend P (1994) Widening inequality of health in northern England, 1981–91. *BMJ* 308:1125–8
- Rothman K, Greenland S (2005) Causation and causal inference in epidemiology. *Am J Public Health* 95(Suppl 1):S144–50
- Seminara NM, Abuabara K, Shin DB *et al.* (2011) Validity of The Health Improvement Network (THIN) for the study of psoriasis. *Br J Dermatol* 164:602–9
- Shah R, Lu Y, Hinkle CC *et al.* (2009) Gene profiling of human adipose tissue during evoked inflammation *in vivo*. *Diabetes* 58:2211–9
- Zimmet P, Magliano D, Matsuzawa Y *et al.* (2005) The metabolic syndrome: a global public health problem and a new definition. *J Atheroscler Thromb* 12:295–300

Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity

To gain further insight into the genetic architecture of psoriasis, we conducted a meta-analysis of 3 genome-wide association studies (GWAS) and 2 independent data sets genotyped on the Immunochip, including 10,588 cases and 22,806 controls. We identified 15 new susceptibility loci, increasing to 36 the number associated with psoriasis in European individuals. We also identified, using conditional analyses, five independent signals within previously known loci. The newly identified loci shared with other autoimmune diseases include candidate genes with roles in regulating T-cell function (such as *RUNX3*, *TAGAP* and *STAT3*). Notably, they included candidate genes whose products are involved in innate host defense, including interferon-mediated antiviral responses (*DDX58*), macrophage activation (*ZC3H12C*) and nuclear factor (NF)- κ B signaling (*CARD14* and *CARM1*). These results portend a better understanding of shared and distinctive genetic determinants of immune-mediated inflammatory disorders and emphasize the importance of the skin in innate and acquired host defense.

Psoriasis is a chronic, potentially disfiguring immune-mediated inflammatory disease of the skin with a prevalence of 0.2–2%, depending on the population of origin. Approximately one-quarter of individuals with psoriasis develop painful and debilitating arthritis, and there is increasing awareness of comorbidities, including metabolic syndromes and cardiovascular disease^{1,2}. Current evidence suggests that a dysregulated cutaneous immune response, characterized by tumor necrosis factor (TNF)- α dependence and exaggerated helper T cell 1 (T_H1) and 17 (T_H17) activation, occurs in genetically susceptible individuals^{1,2}. Recent large-scale association studies have identified 26 loci that are associated with psoriasis^{3–10}, 21 of which show association in individuals of European ancestry^{3–6,10}. Several of these loci overlap with those identified as being associated with other autoimmune diseases (for example, Crohn's disease, ankylosing spondylitis and celiac disease), particularly those near genes involved in T_H17 differentiation and interleukin (IL)-17 responsiveness (for example, *IL23R*, *IL12B*, *IL23A* and *TRAF3IP2*)¹¹. To accelerate the understanding of the genetic architecture of psoriasis, we helped design a custom SNP array—the Immunochip—with the aim of fine mapping previously identified genome-wide significant susceptibility loci (associated at $P < 5 \times 10^{-8}$) and exploring replication of thousands of SNPs representing additional promising signals^{12,13}. In this study, we used Immunochip data to identify new susceptibility loci for psoriasis and to consider their association to other autoimmune disorders.

We combined three existing GWAS data sets (referred to as Kiel³, the Collaborative Association Study of Psoriasis (CASP)⁴ and the Wellcome Trust Case Control Consortium 2 (WTCCC2)⁵) with two independent case-control data sets of individuals of European descent genotyped on the Immunochip: the Psoriasis Association Genetics Extension (PAGE: 3,580 cases and 5,902 controls) and the Genetic Analysis of Psoriasis Consortium (GAPC: 2,997 cases and 9,183 controls) (data sets are described in **Supplementary Tables 1 and 2**). After quality control, the combined data set consisted of 10,588 individuals with psoriasis and 22,806 healthy controls. For each GWAS, we increased the SNP density through imputation, using European haplotype sequences generated by the 1000 Genomes Project (20100804 release) as templates. Overall, our analysis included 111,236 SNPs that were genotyped in both Immunochip data sets that also had good imputation quality in at least 2 of the 3 GWAS (Online Methods).

From meta-analysis of all five data sets, we confirmed associations at genome-wide significance for SNPs at 19 of the 21 known loci involved in psoriasis (**Table 1**, **Supplementary Fig. 1** and **Supplementary Table 3**). We found nominal evidence of association for the remaining two loci in the combined analysis (*ZMIZ1* and *PRDX5*, each with $P < 3 \times 10^{-6}$), as well as nominal evidence of association for all loci in separate analyses including only GWAS (all with $P < 5 \times 10^{-3}$) or Immunochip data (all with $P < 4 \times 10^{-4}$). In addition, we identified 15 new risk loci associated at $P < 5 \times 10^{-8}$ (**Table 1**, **Supplementary Fig. 1** and **Supplementary Table 3**). Nine of the new signals were submitted, during design of the Immunochip, as genome-wide significant Immunochip loci by at least one other disease consortium (**Supplementary Table 4**, disease overlap), although we also submitted three of these (rs11121129, rs10865331 and rs9504361) on the basis of a preliminary meta-analysis of our GWAS data sets. Notably, of the remaining six signals, four were submitted for psoriasis (rs11795343, rs4561177, rs11652075 and rs545979). The strongest new association was observed for rs892085 at 19p13.2 near the *ILF3* and *CARM1* genes (combined P value (P_{combined}) = 3.0×10^{-17} ; odds ratio (OR) = 1.17). Despite its proximity (< 500 kb away) to *TYK2*, conditional analysis showed that this SNP represents an independent signal (**Supplementary Table 5**). Other associated loci included 1p36.11 near *RUNX3*, 6p25.3 near *EXOC2* and *IRF4*, 9p21.1 near *DDX58*, 11q22.3 near *ZC3H12C*, 11q24.3 in the *ETS1* gene and 17q21.2 near *STAT3*, *STAT5A* and *STAT5B*. The functional characteristics of notable genes from the newly identified loci are summarized in **Box 1**, and regional association plots are shown in **Supplementary Figure 2**.

A full list of authors and affiliations appears at the end of the paper.

Received 7 May; accepted 17 October; published online 11 November 2012; doi:10.1038/ng.2467

Table 1 Meta-analysis results for loci associated with psoriasis in a weighted analysis

SNP	Chr.	Position (bp)	GWAS <i>P</i> value (meta)	Immunochip <i>P</i> value (meta)	Combined <i>P</i> value	Risk/non-risk allele	RAF (case)	RAF (control)	OR ^a (meta)	Notable genes	Number of genes ± 500 kb
Known susceptibility loci											
rs7552167	1	24518643	2.3×10^{-5}	8.4×10^{-8}	8.5×10^{-12}	G/A	0.878	0.858	1.21	<i>IL28RA</i>	26
rs9988642	1	67726104	2.5×10^{-13}	3.5×10^{-15}	1.1×10^{-26}	T/C	0.952	0.929	1.52	<i>IL23R</i>	17
rs6677595	1	152590187	8.1×10^{-15}	2.7×10^{-20}	2.1×10^{-33}	T/C	0.689	0.640	1.26	<i>LCE3B, LCE3D</i>	43
rs62149416	2	61083506	3.4×10^{-10}	3.2×10^{-9}	1.8×10^{-17}	T/C	0.671	0.635	1.17	<i>FLJ16341, REL</i>	9
rs17716942	2	163260691	4.1×10^{-9}	1.0×10^{-10}	3.3×10^{-18}	T/C	0.891	0.863	1.27	<i>KCNH7, IFIH1</i>	7
rs27432	5	96119273	4.4×10^{-8}	7.5×10^{-14}	1.9×10^{-20}	A/G	0.309	0.274	1.20	<i>ERAP1</i>	7
rs1295685	5	131996445	8.5×10^{-6}	6.7×10^{-6}	3.4×10^{-10}	G/A	0.807	0.798	1.18	<i>IL13, IL4</i>	21
rs2233278	5	150467189	4.9×10^{-17}	5.2×10^{-27}	2.2×10^{-42}	C/G	0.090	0.058	1.59	<i>TNIP1</i>	17
rs12188300	5	158829527	7.5×10^{-23}	3.3×10^{-32}	3.2×10^{-53}	T/A	0.132	0.095	1.58	<i>IL12B</i>	5
rs4406273	6	31266090	5.3×10^{-300}	3.6×10^{-427}	4.5×10^{-723}	A/G	0.259	0.092	4.32	<i>HLA-B, HLA-C</i>	56
rs33980500	6	111913262	4.3×10^{-20}	7.6×10^{-27}	4.2×10^{-45}	T/C	0.108	0.074	1.52	<i>TRAF3IP2</i>	8
rs582757	6	138197824	2.0×10^{-14}	3.7×10^{-13}	2.2×10^{-25}	C/T	0.315	0.273	1.23	<i>TNFAIP3</i>	5
rs1250546	10	81032532	5.1×10^{-4}	3.2×10^{-4}	6.8×10^{-7}	A/G	0.605	0.579	1.10	<i>ZMIZ1</i>	9
rs645078	11	64135298	4.7×10^{-3}	1.5×10^{-4}	2.2×10^{-6}	A/C	0.626	0.609	1.09	<i>RPS6KA4, PRDX5</i>	36
rs2066819	12	56750204	7.5×10^{-12}	8.9×10^{-8}	5.4×10^{-17}	C/T	0.948	0.934	1.39	<i>STAT2, IL23A</i>	40
rs8016947	14	35832666	1.4×10^{-9}	1.6×10^{-9}	2.5×10^{-17}	G/T	0.600	0.564	1.16	<i>NFKBIA</i>	11
rs12445568	16	31004812	1.2×10^{-6}	1.8×10^{-11}	1.2×10^{-16}	C/T	0.403	0.368	1.16	<i>PRSS53, FBXL19</i>	46
rs28998802	17	26124908	3.6×10^{-6}	1.7×10^{-11}	3.3×10^{-16}	A/G	0.170	0.145	1.22	<i>NOS2</i>	9
rs34536443	19	10463118	5.1×10^{-10}	2.6×10^{-22}	9.1×10^{-31}	G/C	0.974	0.953	1.88	<i>TYK2</i>	42
rs1056198	20	48556229	6.2×10^{-9}	1.6×10^{-7}	1.5×10^{-14}	C/T	0.600	0.573	1.16	<i>RNF114</i>	11
rs4821124	22	21979289	5.4×10^{-5}	1.2×10^{-4}	3.8×10^{-8}	C/T	0.208	0.189	1.13	<i>UBE2L3</i>	16
Newly identified susceptibility loci											
rs11121129	1	8268095	7.3×10^{-5}	4.6×10^{-5}	1.7×10^{-8}	A/G	0.308	0.287	1.13	<i>SLC45A1, TNFRSF9</i>	15
rs7536201	1	25293084	7.8×10^{-5}	6.4×10^{-9}	2.3×10^{-12}	C/T	0.528	0.494	1.13	<i>RUNX3</i>	18
rs10865331	2	62551472	4.5×10^{-4}	2.6×10^{-7}	4.7×10^{-10}	A/G	0.404	0.374	1.12	<i>B3GNT2</i>	6
rs9504361	6	577820	5.1×10^{-7}	4.2×10^{-6}	2.1×10^{-11}	A/G	0.574	0.546	1.12	<i>EXOC2, IRF4</i>	5
rs2451258	6	159506600	4.4×10^{-4}	2.0×10^{-5}	3.4×10^{-8}	C/T	0.362	0.348	1.12	<i>TAGAP</i>	8
rs2700987	7	37386237	3.3×10^{-7}	4.6×10^{-4}	4.3×10^{-9}	A/C	0.591	0.564	1.11	<i>ELMO1</i>	3
rs11795343	9	32523737	2.8×10^{-7}	2.1×10^{-5}	8.4×10^{-11}	T/C	0.628	0.597	1.11	<i>DDX58</i>	7
rs10979182	9	110817020	2.8×10^{-5}	1.2×10^{-4}	2.3×10^{-8}	A/G	0.617	0.591	1.12	<i>KLF4</i>	0
rs4561177	11	109962432	1.1×10^{-4}	1.4×10^{-9}	7.7×10^{-13}	A/G	0.617	0.581	1.14	<i>ZC3H12C</i>	4
rs3802826	11	128406438	1.1×10^{-3}	2.0×10^{-7}	9.5×10^{-10}	A/G	0.505	0.484	1.12	<i>ETS1</i>	7
rs367569	16	11365500	2.6×10^{-4}	4.6×10^{-5}	4.9×10^{-8}	C/T	0.729	0.709	1.13	<i>PRM3, SOCS1</i>	14
rs963986	17	40561579	9.9×10^{-5}	1.2×10^{-5}	5.3×10^{-9}	C/G	0.169	0.154	1.15	<i>PTRF, STAT3, STAT5A/B</i>	42
rs11652075	17	78178893	1.3×10^{-3}	7.0×10^{-6}	3.4×10^{-8}	C/T	0.530	0.502	1.11	<i>CARD14</i>	16
rs545979	18	51819750	1.4×10^{-6}	2.4×10^{-5}	3.5×10^{-10}	T/C	0.317	0.291	1.12	<i>POL1, STARD6, MBD2</i>	6
rs892085	19	10818092	1.2×10^{-7}	4.5×10^{-11}	3.0×10^{-17}	A/G	0.593	0.558	1.17	<i>ILF3, CARM1</i>	37

For previously identified susceptibility loci, the most significantly associated SNP within 500 kb (3 Mb for the MHC region) of the previously published SNP is shown. rs34536443 was the most significantly associated SNP in the *TYK2* region but was found to be independent of the previously published SNP (rs12720356). GWAS *P* value, *P* value from the meta-analysis of the three GWAS data sets; Immunochip *P* value, the result of the meta-analysis of the two Immunochip data sets; combined *P* value, the *P* value from the meta-analysis including all five data sets; RAF, risk allele frequency; notable genes, genes most likely to have an effect on the development of psoriasis; chr., chromosome.

^aThe overall OR was calculated using the effective sample size.

To identify independent secondary signals, we performed conditional analysis, using as covariates the strongest signals from the 34 loci achieving genome-wide significance in this study. We identified secondary signals in five loci: 2q24.2, 5q15, 5q33.3, 6p21.33 and 19q13.2 (**Supplementary Figs. 3 and 4** and **Supplementary Tables 6 and 7**). The strongest signal from the conditional analysis mapped to the major histocompatibility (MHC) region near the *MICA* gene (rs13437088: $P = 3.1 \times 10^{-40}$; OR = 1.32), in agreement with a previous conditional analysis¹⁴. The conditional signal at 5q15 was in the *ERAP2* gene (rs2910686: $P = 2.0 \times 10^{-8}$), which did not show any evidence of association in the unconditioned analysis ($P = 0.46$). Further investigation showed that the risk-increasing alleles at *ERAP1* and the risk-decreasing alleles at *ERAP2* preferentially appear on the same haplotype, and the signal near *ERAP2* was thus masked by *ERAP1* before conditional analysis (**Supplementary Note**).

The strongest conditional signal in the 19q13.2 region was rs12720356 in the *TYK2* gene (OR = 1.25; minor allele frequency (MAF) in controls = 0.09; $P = 3.2 \times 10^{-10}$). The association of this SNP with psoriasis has been previously reported⁵ and is independent of the strongest signal in *TYK2* identified by our meta-analysis (rs34536443: OR = 1.88; MAF in cases = 0.03; $P = 1.5 \times 10^{-39}$). Because rs34536443 was a low-frequency imputed SNP and manifested the greatest effect outside of the MHC region, we directly genotyped this SNP in 3,390 independent samples from Michigan (1,844 cases and 1,546 controls), robustly replicating the association (OR = 2.80; MAF in cases = 0.02; $P = 7.8 \times 10^{-14}$) and experimentally confirming the validity of our imputation procedures.

We next tested for statistical interaction among the top SNPs in the 34 significantly associated loci (**Supplementary Table 8** and **Supplementary Note**). We identified two significant pairwise interactions after

Box 1 Annotated functions of notable candidate genes within newly identified psoriasis susceptibility loci

RRER, SLC45A1, ERFF1 and TNFRSF9 (1p36.23): The signal falls between the *RRER*, *SLC45A1*, *ERFF1* and *TNFRSF9* genes. *RRER* encodes an arginine–glutamic acid dipeptide repeat–containing protein that controls retinoic acid signaling³⁸. *ERFF1* encodes a feedback inhibitor of the epidermal growth factor (EGF) receptor³⁹. *SLC45A1* encodes a solute carrier protein that mediates the uptake of glucose⁴⁰. The *TNFRSF9* gene encodes a costimulatory molecule that has a role in the generation of memory CD8⁺ T cells.

RUNX3 (1p36.11): *RUNX3* encodes a member of the Runt domain–containing family of transcription factors and has an essential role in T-cell biology, particularly in the generation of CD8⁺ T cells. *RUNX3* also has a role in promoting T_H1 differentiation through binding to T-bet⁴¹.

B3GNT2 (2p15): *B3GNT2* is a member of the β -1,3-*N*-acetylglucosaminyltransferase family. It catalyzes the initiation and elongation of poly-*N*-acetylglucosamine chains⁴². Deficiency for it has been shown to result in hyperactivation of lymphocytes⁴³.

EXOC2 and IRF4 (6p25.3): *EXOC2* encodes a component of the multiprotein complex that mediates the docking of exocytic vesicles to the plasma membrane⁴⁴. *IRF4* encodes a transcription factor that regulates *IL17A* promoter activity and controls ROR γ t-dependent T_H17-mediated colitis *in vivo*^{45,46}. *IRF4* also has a role in the stabilization of the T_H17 phenotype through IL-21 (ref. 47) and may regulate CD4 and/or CD8 T-cell differentiation through regulation of *RUNX3* expression⁴⁸.

TAGAP (6q25.3): *TAGAP* encodes a RhoGTPase-activating protein that is involved in T-cell activation⁴⁹.

ELMO1 (7p14.2-7p14.1): *ELMO1* encodes a member of the engulfment and cell motility protein family, which binds to DOCK2 and is essential for Toll-like receptor (TLR7 and TLR9)-mediated interferon (IFN)- α induction by plasmacytoid dendritic cells⁵⁰ and plasmacytoid dendritic cell migration⁵¹. DOCK2 also has a role in antigen uptake and presentation and lymphocyte trafficking⁵¹.

DDX58 (9p21.1): *DDX58* encodes the RIG-I innate antiviral receptor, which recognizes cytosolic double-stranded RNA⁵². It is induced by IFN- γ (ref. 53) and regulates the production of type I and II IFN⁵⁴.

KLF4 (9p31.2): *KLF4* encodes a Kruppel-like transcription factor, which is required for the establishment of skin barrier function⁵⁵ and regulates key signaling pathways related to macrophage activation⁵⁶. The KLF4 protein also binds to the promoter of *IL17A* and positively regulates its expression.

ZC3H12C (11q22.3): *ZC3H12C* encodes a zinc-finger protein regulating macrophage activation⁵⁷.

ETS1 (11q24.3): *ETS1* encodes a transcription factor activated downstream of the Ras–mitogen-activated protein kinase (MAPK) pathway and is involved in the homeostasis of squamous epithelia⁵⁸. It is involved in CD8 T-cell differentiation and acts, in part, by promoting *RUNX3* expression⁵⁹. It is also a negative regulator of T_H17 differentiation⁶⁰.

SOCS1 (16p13.13): *SOCS1* is a member of the suppressor of cytokine signaling family of proteins and inhibits signaling events downstream of IFN- γ (ref. 61). It regulates T_H17 differentiation by maintaining STAT3 transcriptional activity⁶² and interacts with TYK2 in cytokine signaling⁶³.

STAT3, STAT5A and STAT5B (17q21.2): *STAT3*, *STAT5A* and *STAT5B* encode members of the STAT family of transcriptional activators. STAT3 participates in signaling downstream of multiple cytokines implicated in psoriasis, such as IL-6, IL-10, IL-20, IL-22 and IL-23, and may have a role in mediating the innate immune response in psoriatic epidermis⁶⁴. STAT3 is required for the differentiation of T_H17 cells⁶⁵. STAT5A and STAT5B participate in signaling downstream of the IL-2 family of cytokines, including IL-2, IL-7, IL-15 and IL-21. Both proteins contribute to the development of regulatory T (Treg) cells and inhibit the differentiation of T_H17 cells⁶⁶.

CARD14 (17q25.3): *CARD14* encodes a member of the family of caspase recruitment domain–containing scaffold proteins, known as CARD- and membrane-associated guanylate kinase–like domain–containing protein (CARMA). *CARD14* (also known as CARMA2) is primarily expressed in epithelial tissues and mediates recruitment and activation of the NF- κ B pathway⁶⁷.

MBD2, POLI and STARD6 (18q21.2): *MBD2* encodes a transcriptional repressor that binds to methylated DNA and has a role in the generation of memory CD8⁺ T cells⁶⁸. *POLI* encodes an error-prone DNA polymerase, which contributes to the hypermutation of immunoglobulin genes⁶⁹. Sterol transport is mediated by vesicles or by soluble protein carriers such as steroidogenic acute regulatory protein (STAR; MIM 600617). STAR is homologous to a family of proteins containing a STAR-related lipid transfer (START) domain of 200–210 amino acids in length, including STARD6.

ILF3 and CARM1 (19p13.2): *ILF3* encodes a double-stranded RNA-binding protein that complexes with other proteins, double-stranded RNAs, small noncoding RNAs and mRNAs to regulate gene expression and stabilize mRNA. It is a subunit of the nuclear factor of activated T cells (NFAT), a transcription factor required for the expression in T cells of *IL-2*. *CARM1* encodes a transcriptional coactivator of NF- κ B and functions as a promoter-specific regulator of NF- κ B recruitment to chromatin.

correction for multiple testing ($P < 5 \times 10^{-5}$): *HLA-C* (rs4406273) with *LCE* (rs6677595) and *HLA-C* (rs4406273) with *ERAP1* (rs27432). These interactions confirm the results of previous studies^{5,15,16}.

To identify potential causal alleles in coding sequence, we looked for missense variants in tight linkage disequilibrium (LD; $r^2 > 0.9$ in 1000 Genomes Project European samples) with the lead SNP at each of the 34 identified loci (Table 1 and Supplementary Table 6). We found ten potentially causal SNPs (Table 2), nine of which were included in our meta-analysis. For the known loci near *TRAF3IP2* and *TYK2*, damaging nonsynonymous substitutions were themselves the index SNPs in our initial and conditional analyses.

For the newly identified loci, the index SNP from *CARD14*, a gene that harbors Mendelian variants predisposing to psoriasis¹⁷, was also a common and damaging variant, as has been described elsewhere¹⁸. For the remaining loci, we could account for essentially all index SNP signals by conditioning on nearby missense SNPs, consistent with the possibility that they are causal. Notable nonsynonymous variants included the protective polymorphism encoding a p.Arg381Gln alteration in *IL23R*¹⁹, a SNP in the *PRSS53* gene²⁰, which is also the most highly overexpressed gene in psoriatic skin in this locus⁶, and a variant in *YDJC* that also increases risk for celiac disease²¹, rheumatoid arthritis²² and Crohn's disease²³.

Table 2 Potential causal SNPs for significant loci

Index SNP ^a			Potential causal SNP ^b				Combined <i>P</i> value ^c			
Marker ^d	RAF	Annotation	Marker ^f	RAF	Gene with variant	Amino-acid substitution (damaging effect) ^g	<i>r</i> ²	Index SNP	Potential causal SNP	Index SNP (conditioning on causal SNP)
rs9988642	0.93	454 bp downstream of <i>IL23R</i>	rs11209026	0.94	<i>IL23R</i>	p.Arg381Gln (P)	0.91	1.1 × 10 ⁻²⁶	1.5 × 10 ⁻²⁶	0.13
rs27432	0.29	Intron <i>ERAP1</i>	rs27044	0.29	<i>ERAP1</i>	p.Gln730Glu	1	1.9 × 10 ⁻²⁰	2.3 × 10 ⁻²⁰	0.14
rs1295685	0.77	3' UTR <i>IL13</i>	rs20541	0.77	<i>IL13</i>	p.Arg144Gln	0.97	3.4 × 10 ⁻¹⁰	3.5 × 10 ⁻¹⁰	0.78
rs33980500	0.09	Missense	Self	0.09	<i>TRAF3IP2</i>	p.Asp19Asn (S,P)	1	4.2 × 10 ⁻⁴⁵	4.2 × 10 ⁻⁴⁵	NA
rs2066819	0.93	Intron <i>STAT2</i>	rs2066807	0.93	<i>STAT2</i>	p.Met594Ile	0.9	5.4 × 10 ⁻¹⁷	5.1 × 10 ⁻¹⁶	0.036
rs12445568	0.36	Intron <i>STX1B</i>	rs7199949	0.37	<i>PRSS53</i>	p.Pro406Ala	0.9	1.2 × 10 ⁻¹⁶	—	—
rs11652075	0.51	Missense	Self	0.51	<i>CARD14</i>	p.Arg820Trp (S)	1	3.4 × 10 ⁻⁸	3.4 × 10 ⁻⁸	NA
rs34536443	0.97	Missense	Self	0.97	<i>TYK2</i>	p.Pro1104Ala (S,P)	1	1.5 × 10 ⁻³⁹	1.5 × 10 ⁻³⁹	NA
rs12720356 ^e	0.9	Missense	Self	0.9	<i>TYK2</i>	p.Ile684Ser (S,P)	1	3.2 × 10 ⁻¹⁰	3.2 × 10 ⁻¹⁰	NA
rs4821124	0.19	966 bp downstream of <i>UBE2L3</i>	rs2298428	0.18	<i>YDJC</i>	p.Ala263Thr	0.96	3.8 × 10 ⁻⁸	6.2 × 10 ⁻⁸	0.48

NA, not available, as the index SNP is the potential causal SNP. Shown are data for SNPs that are missense mutations in 1000 Genome Project data that were in LD ($r^2 \geq 0.9$) with primary signals from the known and newly identified loci that achieved genome-wide significance in the meta-analysis or with secondary signals from the conditional analysis (Index SNP).

^aInformation for SNPs with the most significant *P* value in our analysis. ^bInformation for the SNPs that had high LD with the strongest signal. ^cMeta-analysis *P* value for the index SNP and potential causal SNP and the *P* value for the index SNP while conditioning on the potential causal SNP. Note that the potential causal SNP rs7199949 is not present in our meta-analysis study; therefore its *P* value was not shown. ^dSNPs with the most significant *P* value in our analysis. ^eMeta-analysis *P* value from the conditional analysis. ^fSNPs that are missense mutations that had high LD with our strongest signal. ^gHigh-confidence damaging effect predicted by SIFT (S) or Polyphen (P).

Using the results of a large-scale study of gene expression in psoriatic versus normal skin²⁴, we found 14 upregulated genes (*IL12RB2*, *LCE3D*, *REL*, *PUS10*, *CDSN*, *PRSS53*, *PRSS8*, *NOS2*, *DDX58*, *ZC3H12C*, *SOC1*, *STAT3*, *CARD14* and *IFIH1*) and 4 downregulated genes (*MICA*, *RNF114*, *PTRF* and *POLI*) in the 34 associated regions (false discovery rate (FDR) of <0.05; fold change of >1.5 or <0.67; **Supplementary Table 9**). The number of differentially expressed genes in psoriasis susceptibility loci was not greater than expected by chance ($P = 0.39$). None of the 34 top SNPs met the Bonferroni-corrected significance threshold ($P < 1 \times 10^{-7}$) for expression quantitative trait loci (eQTLs) in skin tissue, as assessed by microarray analysis of mRNA levels²⁵. However, rs2910686, one of the five SNPs identified by conditional analysis, was a *cis* eQTL for *ERAP2* in both normal and psoriatic skin (**Supplementary Note**). Genetic control of *ERAP2* expression has been noted previously^{26,27} and has been suggested as a determinant of balancing selection at this locus²⁸.

This study increases the number of psoriasis-associated regions in individuals of European ancestry to 36, with conditional analysis increasing the number of independent signals to 41. The 39 independent signals associated at $P < 5 \times 10^{-8}$ in the current study collectively account for 14.3% of the total variance in psoriasis risk or approximately 22% of its estimated heritability²⁹ (**Supplementary Table 10**), indicating that further genetic studies, including fine-mapping studies and searches for uncommon susceptibility variants, are in order.

Sharing of susceptibility loci between autoimmune diseases has been demonstrated previously¹¹, and we found similar patterns of overlap in this study. Notably, ten of the psoriasis susceptibility loci reported here overlapped with those reported in Crohn's disease and ten others associated with celiac disease, diseases that occur at higher frequencies in individuals with psoriasis^{30,31} (**Supplementary Fig. 5** and **Supplementary Table 4**). We caution that the statistical significance of these overlapping loci is hard to assess given the ongoing process of gene discovery for many autoimmune disorders and biases in the list of SNPs evaluated for association in this experiment.

As the primary interface with the external environment, the skin provides a critical first line of host defense to microbial pathogens. Consistent with this function, it possesses a diverse and well-conserved set of innate immune mechanisms^{32,33} that emerged long before the development of adaptive immunity³⁴. In this context, we found it noteworthy that five of the six newly identified loci that are thus far uniquely associated with psoriasis are involved in innate

immune responses (*DDX58*, *KLF4*, *ZC3H12C*, *CARD14* and *CARM1*; **Box 1** and **Supplementary Table 4**). Among all confirmed psoriasis susceptibility loci, 11 out of 14 psoriasis-specific loci (the 5 new loci involved in the innate immune response along with *IL28RA*, *LCE3D*, *NOS2*, *FBXL19*, *NFKB1A* and *RNF114*) encode plausible regulators of innate host defense^{1,2,35}. Conversely, only 6 out of 20 loci shared with other autoimmune diseases contain genes that contribute to innate immunity (*REL*, *IFIH1*, *TNIP1*, *TNFAIP3*, *IRF4* and *ELMO1*). These provisional comparisons further illustrate the insights that can be gained by developing and comparing complete and well-annotated sets of risk loci for autoimmune disorders.

The known and newly identified psoriasis susceptibility loci implicated by this study encode several proteins engaged in the TNF, IL-23 and IL-17 signaling pathways, which are targeted by highly effective biological therapies³⁶. Notably, our strongest non-MHC signal directly implicates *TYK2*, a druggable target that contributes to several autoimmune diseases. Agents targeting the closely related JAK kinases are showing encouraging results in clinical trials³⁷. Our findings will help prioritize and interpret the results of sequencing and gene expression studies. Further genomic studies are needed to identify the underlying causal variants in these psoriasis susceptibility loci and to bring increased understanding of pathogenetic mechanisms.

URLs. WTCCC common controls, <http://www.wtccc.org.uk/>; 1000 Genomes Project data, <ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20100804/>; Catalog of Published Genome-Wide Association Studies, <http://www.genome.gov/gwastudies/>; eQTL database, <http://www.sph.umich.edu/csg/jundung/eQTL/TableDownload/>.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

Major support for this study was provided by the US National Institutes of Health, the Wellcome Trust and the German Research Foundation. We thank J.C. Barrett for contribution to the design of the ImmunoChip and helpful analytical discussion, as well as E. Gray, S. Bumpstead, D. Simpkin and the staff of the Wellcome Trust Sanger Institute Sample Management and Genotyping teams for their genotyping and analytical contributions. We acknowledge use of the British 1958 Birth Cohort DNA collection, funded by the UK Medical Research Council (G0000934) and the

Wellcome Trust (068545/Z/02), and the UK National Blood Service controls, funded by the Wellcome Trust. We acknowledge CASP for the contribution of GWAS data, as well as the provision of control DNA samples by the Cooperative Research in the Region of Augsburg (KORA) and Heinz-Nixdorf Recall (Risk Factors, Evaluation of Coronary Calcification and Lifestyle) study (HNR) and genotyping data generated by the Dietary, Lifestyle and Genetic determinants of Obesity and Metabolic syndrome (DILGOM) Consortium. We thank the Barbara and Neal Henschel Charitable Foundation for their support of the National Psoriasis Victor Henschel BioBank. We acknowledge the Genetic repository in Ireland for Psoriasis and Psoriatic Arthritis (GRIPPSA), the Irish blood transfusion service/Trinity College Dublin Biobank and the Dublin Centre for Clinical Research (funded by the Health Research Board and the Wellcome Trust). A detailed list of contributing consortia and relevant funding support is provided in the **Supplementary Note**.

AUTHOR CONTRIBUTIONS

J.T.E., R.C.T. and G.R.A. designed and directed the study. R.P.N., M.W., J.D., J.J.V., J.T.E., F.C., J.N.W.N.B., M.H.A., C.H.S., A.D.B., C.E.M.G., A.R., J. Kere, X.E., W.W., J. Worthington, R.T.-A., M.S., G.N., L.S., R.M., M.J.C., J.S., A.F., S.W., S.K., K.K., T.E., A.M., A.M.B., G.G.K., D.D.G., P.R., U.M., F.O.N., A.H., J. Winkelmann, S.S., C.W., C.L., S.E., R.A., V.C., C.F.R., H.B., H.W.L. and H.E.W. contributed to sample collection and phenotyping. J. Knight coordinated the samples and data sets for the Genetic Analysis of Psoriasis (GAP) consortium. J.T.E. coordinated the PAGE samples and data sets. P.D., A.S., G.B., R.D.P., D.V. and C.C.A.S. contributed to the design of the Immunochip. J. Knight, P.E.S., G.R.A. and H.M.K. advised on the statistical analysis. C.L., S.E., R.A., H.B., E.E., P.H. and R.P.N. performed genotyping. E.E., S.L.S., L.C.T. and H.M.K. performed the genotype calling. S.L.S., L.C.T., Y.L. and J.D. performed genotype imputation and statistical analysis. F.C., J.N.W.N.B., J.E.G., T.T., J.T.E. and A.F. prepared **Box 1**. L.C.T., S.L.S., F.C. and J.T.E. drafted the manuscript and prepared the figures and tables. E.E., J.E.G., J. Knight, P.E.S., R.P.N., R.C.T., T.T., G.R.A., J.N.W.N.B. and A.F. edited and revised the manuscript. All authors approved the final draft.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/doi/10.1038/ng.2467>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Nestle, F.O., Kaplan, D.H. & Barker, J. Psoriasis. *N. Engl. J. Med.* **361**, 496–509 (2009).
- Elder, J.T. *et al.* Molecular dissection of psoriasis: integrating genetics and biology. *J. Invest. Dermatol.* **130**, 1213–1226 (2010).
- Ellinghaus, E. *et al.* Genome-wide association study identifies a psoriasis susceptibility locus at *TRAF3IP2*. *Nat. Genet.* **42**, 991–995 (2010).
- Nair, R.P. *et al.* Genome-wide scan reveals association of psoriasis with IL-23 and NF- κ B pathways. *Nat. Genet.* **41**, 199–204 (2009).
- Strange, A. *et al.* A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between *HLA-C* and *ERAP1*. *Nat. Genet.* **42**, 985–990 (2010).
- Stuart, P.E. *et al.* Genome-wide association analysis identifies three psoriasis susceptibility loci. *Nat. Genet.* **42**, 1000–1004 (2010).
- Sun, L.D. *et al.* Association analyses identify six new psoriasis susceptibility loci in the Chinese population. *Nat. Genet.* **42**, 1005–1009 (2010).
- Zhang, X.J. *et al.* Psoriasis genome-wide association study identifies susceptibility variants within *LCE* gene cluster at 1q21. *Nat. Genet.* **41**, 205–210 (2009).
- de Cid, R. *et al.* Deletion of the late cornified envelope *LCE3B* and *LCE3C* genes as a susceptibility factor for psoriasis. *Nat. Genet.* **41**, 211–215 (2009).
- Ellinghaus, D. *et al.* Combined analysis of genome-wide association studies for Crohn disease and psoriasis identifies seven shared susceptibility loci. *Am. J. Hum. Genet.* **90**, 636–647 (2012).
- Cotsapas, C. *et al.* Pervasive sharing of genetic effects in autoimmune disease. *PLoS Genet.* **7**, e1002254 (2011).
- Cortes, A. & Brown, M.A. Promise and pitfalls of the Immunochip. *Arthritis Res. Ther.* **13**, 101 (2011).
- Trynka, G. *et al.* Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat. Genet.* **43**, 1193–1201 (2011).
- Feng, B.J. *et al.* Multiple loci within the major histocompatibility complex confer risk of psoriasis. *PLoS Genet.* **5**, e1000606 (2009).
- Zheng, H.F. *et al.* Variants in *MHC*, *LCE* and *IL12B* have epistatic effects on psoriasis risk in Chinese population. *J. Dermatol. Sci.* **61**, 124–128 (2011).
- Riveira-Munoz, E. *et al.* Meta-analysis confirms the *LCE3C*, *LCE3B* deletion as a risk factor for psoriasis in several ethnic groups and finds interaction with *HLA-Cw6*. *J. Invest. Dermatol.* **131**, 1105–1109 (2011).
- Jordan, C.T. *et al.* PSORS2 is due to mutations in *CARD14*. *Am. J. Hum. Genet.* **90**, 784–795 (2012).
- Jordan, C.T. *et al.* Rare and common variants in *CARD14*, encoding an epidermal regulator of NF- κ B, in psoriasis. *Am. J. Hum. Genet.* **90**, 796–808 (2012).
- Di Meglio, P. *et al.* The IL23R R381Q gene variant protects against immune-mediated diseases by impairing IL-23-induced Th17 effector response in humans. *PLoS ONE* **6**, e17160 (2011).
- Cal, S. *et al.* Identification and characterization of human polyserase-3, a novel protein with tandem serine-protease domains in the same polypeptide chain. *BMC Biochem.* **7**, 9 (2006).
- Dubois, P.C. *et al.* Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* **42**, 295–302 (2010).
- Zhernakova, A. *et al.* Meta-analysis of genome-wide association studies in celiac disease and rheumatoid arthritis identifies fourteen non-HLA shared loci. *PLoS Genet.* **7**, e1002004 (2011).
- Franke, A. *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.* **42**, 1118–1125 (2010).
- Gudjonsson, J.E. *et al.* Assessment of the psoriatic transcriptome in a large sample: additional regulated genes and comparisons with *in vitro* models. *J. Invest. Dermatol.* **130**, 1829–1840 (2010).
- Ding, J. *et al.* Gene expression in skin and lymphoblastoid cells: refined statistical method reveals extensive overlap in *cis*-eQTL signals. *Am. J. Hum. Genet.* **87**, 779–789 (2010).
- Mason, C.C. *et al.* Bimodal distribution of RNA expression levels in human skeletal muscle tissue. *BMC Genomics* **12**, 98 (2011).
- Song, M.Y., Kim, H.E., Kim, S., Choi, I.H. & Lee, J.K. SNP-based large-scale identification of allele-specific gene expression in human B cells. *Gene* **493**, 211–218 (2012).
- Andrés, A.M. *et al.* Balancing selection maintains a form of *ERAP2* that undergoes nonsense-mediated decay and affects antigen presentation. *PLoS Genet.* **6**, e1001157 (2010).
- Grijbovski, A.M., Olsen, A.O., Magnus, P. & Harris, J.R. Psoriasis in Norwegian twins: contribution of genetic and environmental effects. *J. Eur. Acad. Dermatol. Venerol.* **21**, 1337–1343 (2007).
- Najarian, D.J. & Gottlieb, A.B. Connections between psoriasis and Crohn's disease. *J. Am. Acad. Dermatol.* **48**, 805–821 quiz 822–824 (2003).
- Ludvigsson, J.F., Lindelof, B., Zingone, F. & Ciacci, C. Psoriasis in a nationwide cohort study of patients with celiac disease. *J. Invest. Dermatol.* **131**, 2010–2016 (2011).
- Modlin, R.L. Innate immunity: ignored for decades, but not forgotten. *J. Invest. Dermatol.* **132**, 882–886 (2012).
- Nakatsui, T. & Gallo, R.L. Antimicrobial peptides: old molecules with new ideas. *J. Invest. Dermatol.* **132**, 887–895 (2012).
- Wölfe, U., Martin, S., Emde, M. & Schempp, C. Dermatology in the Darwin anniversary. Part 2: Evolution of the skin-associated immune system. *J. Dtsch. Dermatol. Ges.* **7**, 862–869 (2009).
- Capon, F., Burden, A.D., Trembath, R.C. & Barker, J.N. Psoriasis and other complex trait dermatoses: from loci to functional pathways. *J. Invest. Dermatol.* **132**, 915–922 (2012).
- Gudjonsson, J.E. & Elder, J.T. Psoriasis. in *Dermatology in General Medicine* Vol. 1 (eds. Goldsmith, L. *et al.*) 197–231 (McGraw-Hill, New York, 2012).
- Garber, K. Psoriasis: from bed to bench and back. *Nat. Biotechnol.* **29**, 563–566 (2011).
- Vilhais-Neto, G.C. *et al.* Rere controls retinoic acid signalling and somite bilateral symmetry. *Nature* **463**, 953–957 (2010).
- Ferby, I. *et al.* Mig6 is a negative regulator of EGF receptor-mediated skin morphogenesis and tumor formation. *Nat. Med.* **12**, 568–573 (2006).
- Amler, L.C. *et al.* Identification and characterization of novel genes located at the t(1;15)(p36;q24) translocation breakpoint in the neuroblastoma cell line NGP. *Genomics* **64**, 195–202 (2000).
- Djuretic, I.M. *et al.* Transcription factors T-bet and Runx3 cooperate to activate *Ifng* and silence *Il4* in T helper type 1 cells. *Nat. Immunol.* **8**, 145–153 (2007).
- Shiraishi, N. *et al.* Identification and characterization of three novel β 1,3-N-acetylglucosaminyltransferases structurally related to the β 1,3-galactosyltransferase family. *J. Biol. Chem.* **276**, 3498–3507 (2001).
- Togayachi, A. β 3GnT2 (*B3GNT2*), a major polylactosamine synthase: analysis of *B3GNT2*-deficient mice. *Methods Enzymol.* **479**, 185–204 (2010).
- Grindstaff, K.K. *et al.* Sec6/8 complex is recruited to cell-cell contacts and specifies transport vesicle delivery to the basal-lateral membrane in epithelial cells. *Cell* **93**, 731–740 (1998).
- Biswas, P.S. *et al.* Phosphorylation of IRF4 by ROCK2 regulates IL-17 and IL-21 production and the development of autoimmunity in mice. *J. Clin. Invest.* **120**, 3280–3295 (2010).
- Mudter, J. *et al.* IRF4 regulates IL-17A promoter activity and controls ROR γ -dependent Th17 colitis *in vivo*. *Inflamm. Bowel Dis.* **17**, 1343–1358 (2011).
- Huber, M. *et al.* IRF4 is essential for IL-21-mediated induction, amplification, and stabilization of the Th17 phenotype. *Proc. Natl. Acad. Sci. USA* **105**, 20846–20851 (2008).
- Bowcock, A.M. *et al.* Insights into psoriasis and other inflammatory diseases from large-scale gene expression studies. *Hum. Mol. Genet.* **10**, 1793–1805 (2001).
- Chang, I.F. & Hsiao, H.Y. Induction of RhoGAP and pathological changes characteristic of Alzheimer's disease by UAHFEMF discharge in rat brain. *Curr. Alzheimer Res.* **2**, 559–569 (2005).
- Gotoh, K. *et al.* Selective control of type I IFN induction by the Rac activator DOCK2 during TLR-mediated plasmacytoid dendritic cell activation. *J. Exp. Med.* **207**, 721–730 (2010).

51. Ippagunta, S.K. *et al.* The inflammasome adaptor ASC regulates the function of adaptive immune cells by controlling Dock2-mediated Rac activation and actin polymerization. *Nat. Immunol.* **12**, 1010–1016 (2011).
52. Loo, Y.M. & Gale, M. Immune signaling by RIG-I-like receptors. *Immunity* **34**, 680–692 (2011).
53. Cui, X.F., Imaizumi, T., Yoshida, H., Borden, E.C. & Satoh, K. Retinoic acid-inducible gene-I is induced by interferon- γ and regulates the expression of interferon- γ stimulated gene 15 in MCF-7 cells. *Biochem. Cell Biol.* **82**, 401–405 (2004).
54. Negishi, H. *et al.* A critical link between Toll-like receptor 3 and type II interferon signaling pathways in antiviral innate immunity. *Proc. Natl. Acad. Sci. USA* **105**, 20446–20451 (2008).
55. Patel, S., Xi, Z.F., Seo, E.Y., McGaughey, D. & Segre, J.A. Klf4 and corticosteroids activate an overlapping set of transcriptional targets to accelerate *in utero* epidermal barrier acquisition. *Proc. Natl. Acad. Sci. USA* **103**, 18668–18673 (2006).
56. Feinberg, M.W. *et al.* Kruppel-like factor 4 is a mediator of proinflammatory signaling in macrophages. *J. Biol. Chem.* **280**, 38247–38258 (2005).
57. Liang, J. *et al.* A novel CCCH-zinc finger protein family regulates proinflammatory activation of macrophages. *J. Biol. Chem.* **283**, 6337–6346 (2008).
58. Nagarajan, P. *et al.* Ets1 blocks terminal differentiation of keratinocytes and induces expression of matrix metalloproteases and innate immune mediators. *J. Cell Sci.* **123**, 3566–3575 (2010).
59. Zamisch, M. *et al.* The transcription factor Ets1 is important for CD4 repression and Runx3 up-regulation during CD8 T cell differentiation in the thymus. *J. Exp. Med.* **206**, 2685–2699 (2009).
60. Moisan, J., Grenningloh, R., Bettelli, E., Oukka, M. & Ho, I.C. Ets-1 is a negative regulator of Th17 differentiation. *J. Exp. Med.* **204**, 2825–2835 (2007).
61. Sakamoto, H. *et al.* A Janus kinase inhibitor, JAB, is an interferon- γ -inducible gene and confers resistance to interferons. *Blood* **92**, 1668–1676 (1998).
62. Tanaka, K. *et al.* Loss of suppressor of cytokine signaling 1 in helper T cells leads to defective Th17 differentiation by enhancing antagonistic effects of IFN- γ on STAT3 and Smads. *J. Immunol.* **180**, 3746–3756 (2008).
63. Piganis, R.A. *et al.* Suppressor of cytokine signaling (SOCS) 1 inhibits type I interferon (IFN) signaling via the interferon α receptor (IFNAR1)-associated tyrosine kinase Tyk2. *J. Biol. Chem.* **286**, 33811–33818 (2011).
64. Sano, S. *et al.* Stat3 links activated keratinocytes and immunocytes required for development of psoriasis in a novel transgenic mouse model. *Nat. Med.* **11**, 43–49 (2005).
65. Harris, T.J. *et al.* Cutting edge: an *in vivo* requirement for STAT3 signaling in T_H17 development and T_H17-dependent autoimmunity. *J. Immunol.* **179**, 4313–4317 (2007).
66. Wei, L., Laurence, A. & O'Shea, J.J. New insights into the roles of Stat5a/b and Stat3 in T cell development and differentiation. *Semin. Cell Dev. Biol.* **19**, 394–400 (2008).
67. Blonska, M. & Lin, X. NF- κ B signaling pathways regulated by CARMA family of scaffold proteins. *Cell Res.* **21**, 55–70 (2011).
68. Kersh, E.N. Impaired memory CD8 T cell development in the absence of methyl-CpG-binding domain protein 2. *J. Immunol.* **177**, 3821–3826 (2006).
69. Faili, A. *et al.* Induction of somatic hypermutation in immunoglobulin genes is dependent on DNA polymerase ι . *Nature* **419**, 944–947 (2002).

Lam C Tsoi^{1,51}, Sarah L Spain^{2,51}, Jo Knight^{3,4,51}, Eva Ellinghaus^{5,51}, Philip E Stuart⁶, Francesca Capon², Jun Ding¹, Yanming Li¹, Trilokraj Tejasvi⁶, Johann E Gudjonsson⁶, Hyun M Kang¹, Michael H Allen², Ross McManus^{7,8}, Giuseppe Novelli^{9,10}, Lena Samuelsson¹¹, Joost Schalkwijk¹², Mona Ståhle¹³, A David Burden¹⁴, Catherine H Smith¹⁵, Michael J Cork¹⁶, Xavier Estivill¹⁷, Anne M Bowcock¹⁸, Gerald G Krueger¹⁹, Wolfgang Weger²⁰, Jane Worthington²¹, Rachid Tazi-Ahnini¹⁶, Frank O Nestle², Adrian Hayday²², Per Hoffmann^{23,24}, Juliane Winkelmann^{25–27}, Cisca Wijmenga²⁸, Cordelia Langford²⁹, Sarah Edkins²⁹, Robert Andrews²⁹, Hannah Blackburn²⁹, Amy Strange³⁰, Gavin Band³⁰, Richard D Pearson³⁰, Damjan Vukcevic³⁰, Chris C A Spencer³⁰, Panos Deloukas²⁹, Ulrich Mrowietz³¹, Stefan Schreiber^{5,32}, Stephan Weidinger³¹, Sulev Koks³³, Külli Kingo³⁴, Tonu Esko³⁵, Andres Metspalu³⁵, Henry W Lim³⁶, John J Voorhees⁶, Michael Weichenthal³¹, H Erich Wichmann^{37–39}, Vinod Chandran⁴⁰, Cheryl F Rosen⁴¹, Proton Rahman⁴², Dafna D Gladman⁴⁰, Christopher E M Griffiths⁴³, Andre Reis⁴⁴, Juha Kere^{45–47}, Collaborative Association Study of Psoriasis⁴⁸, Genetic Analysis of Psoriasis Consortium⁴⁸, Psoriasis Association Genetics Extension⁴⁸, Wellcome Trust Case Control Consortium²⁴⁸, Rajan P Nair⁶, Andre Franke⁵, Jonathan N W N Barker^{2,15}, Goncalo R Abecasis¹, James T Elder^{6,49} & Richard C Trembath^{2,50}

Collaborative Association Study of Psoriasis (CASP):

Kristina Callis Duffin¹⁹, Cindy Helms¹⁸, David Goldgar¹⁹, Yun Li¹, Justin Paschall⁵², Mary J Malloy^{53,54}, Clive R Pullinger^{53,54}, John P Kane^{53,54}, Jennifer Gardner¹⁸, Amy Perlmutter⁵⁵, Andrew Miner⁵⁵, Bing Jian Feng¹⁹, Ravi Hiremagalore⁶, Robert W Ike⁵⁶, Enno Christophers³¹, Tilo Henseler³¹, Andreas Ruether⁵, Steven J Schrod⁵⁷, Sampath Prahalad^{58–60}, Stephen L Guthery^{58–60}, Judith Fischer⁶¹, Wilson Liao⁶², Pui Kwok⁶², Alan Menter⁶³, G Mark Lathrop⁶¹, C Wise⁶⁴ & Ann B Begovich⁵⁷

Genetic Analysis of Psoriasis Consortium:

Alexandros Onoufriadis², Michael E Weale², Angelika Hofer²⁰, Wolfgang Salmhofer²⁰, Peter Wolf²⁰, Kati Kainu⁶⁵, Ulpu Saarialho-Kere⁶⁵, Sari Suomela⁶⁵, Petra Badorf⁴⁴, Ulrike Hüffmeier⁴⁴, Werner Kurrat⁶⁶, Wolfgang Küster⁶⁷, Jesús Lascorz⁶⁸, Rotraut Mössner⁶⁹, Funda Schürmeier-Horst⁷⁰, Markward Ständer⁷¹, Heiko Traupe⁷⁰, Judith G M Bergboer¹², Martin den Heijer^{72,73}, Peter C van de Kerkhof¹², Patrick L J M Zeeuwen¹², Louise Barnes^{7,8}, Linda E Campbell⁷⁴, Caitriona Cusack⁷⁵, Ciara Coleman^{7,8}, Judith Conroy^{7,8}, Sean Ennis^{7,8}, Oliver Fitzgerald⁷⁶, Phil Gallagher⁷⁶, Alan D Irvine⁷⁷, Brian Kirby⁷⁶, Trevor Markham⁷⁵, W H Irwin McLean⁷⁴, Joe McPartlin^{7,8}, Sarah F Rogers⁷⁶, Anthony W Ryan^{7,8}, Agnieszka Zawirska⁷⁶, Emiliano Giardina⁹, Tiziana Lepre⁹, Carlo Perricone⁹, Gemma Martín-Ezquerro⁷⁸, Ramon M Pujol⁷⁸, Eva Riveira-Munoz¹⁷, Annica Inerot⁷⁹, Åsa T Naluai¹¹, Lotus Mallbris¹³, Katarina Wolk¹³, Joyce Leman¹⁴, Anne Barton²¹, Richard B Warren⁴³, Helen S Young⁴³, Isis Ricano-Ponce²⁸ & Gosia Trynka²⁸

Psoriasis Association Genetics Extension:

Fawnda J Pellett⁴⁰, Andrew Henschel⁸⁰, Marin Aurand⁸⁰, Bruce Bebo⁸⁰, Christian Gieger⁸¹, Thomas Illig⁸², Susanne Moebus⁸³, Karl-Heinz Jöckel⁸³ & Raimund Erbel⁸⁴

Wellcome Trust Case Control Consortium 2:

Peter Donnelly^{30,85}, Leena Peltonen²⁹, Jenefer M Blackwell^{86,87}, Elvira Bramon^{88,89}, Matthew A Brown⁹⁰, Juan P Casas⁹¹, Aiden Corvin⁹², Nicholas Craddock⁹³, Audrey Duncanson⁹⁴, Janusz Jankowski⁹⁵, Hugh S Markus⁹⁶, Christopher G Mathew², Mark I McCarthy⁹⁷, Colin N A Palmer⁹⁸, Robert Plomin⁹⁹, Anna Rautanen³⁰, Stephen J Sawcer¹⁰⁰, Nilesh Samani¹⁰¹, Ananth C Viswanathan^{102,103}, Nicholas W Wood¹⁰⁴, Céline Bellenguez³⁰, Colin Freeman³⁰, Garrett Hellenthal³⁰, Eleni Giannoulidou³⁰, Matti Pirinen³⁰, Zhan Su³⁰, Sarah E Hunt²⁹, Rhian Gwilliam²⁹, Suzannah J Bumpstead²⁹, Serge Dronov²⁹, Matthew Gillman²⁹, Emma Gray²⁹, Naomi Hammond²⁹, Alagurevathi Jayakumar²⁹, Owen T McCann²⁹, Jennifer Liddle²⁹, Marc L Perez²⁹, Simon C Potter²⁹, Radhi Ravindrarajah²⁹, Michelle Ricketts²⁹, Matthew Waller²⁹, Paul Weston²⁹, Sara Widaa²⁹ & Pamela Whittaker²⁹

¹Department of Biostatistics, Center for Statistical Genetics, University of Michigan Ann Arbor, Michigan, USA. ²Division of Genetics and Molecular Medicine, King's College London, London, UK. ³Neuroscience Research, Centre for Addiction and Mental Health, Toronto, Ontario, Canada. ⁴National Institute for Health Research (NIHR), Biomedical Research Centre, Guy's and St. Thomas' National Health Service (NHS) Foundation Trust, London, UK. ⁵Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany. ⁶Department of Dermatology, University of Michigan, Ann Arbor, Michigan, USA. ⁷Department of Clinical Medicine, Trinity College Dublin, St. James's Hospital, Dublin, Ireland. ⁸Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland. ⁹National Agency for Evaluation of Universities and Research Institutes (ANVUR), Rome, Italy. ¹⁰Research Center San Pietro Hospital, Rome, Italy. ¹¹Department of Medical and Clinical Genetics, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. ¹²Department of Dermatology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. ¹³Dermatology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden. ¹⁴Department of Dermatology, Western Infirmary, Glasgow, UK. ¹⁵St. John's Institute of Dermatology, King's College London, London, UK. ¹⁶Academic Unit of Dermatology Research, Department of Infection and Immunity, The University of Sheffield, Sheffield, UK. ¹⁷Genes and Disease Programme, Centre for Genomic Regulation (CRG) and Pompeu Fabra University, Hospital del Mar Research Institute and Public Health and Epidemiology Network Biomedical Research Centre (CIBERESP), Barcelona, Spain. ¹⁸Department of Genetics, Division of Human Genetics, Washington University School of Medicine, St. Louis, Missouri, USA. ¹⁹Department of Dermatology, University of Utah, Salt Lake City, Utah, USA. ²⁰Department of Dermatology, Medical University of Graz, Graz, Austria. ²¹Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. ²²Division of Immunology, Infection and Inflammatory Disease, King's College London, London, UK. ²³Institute of Human Genetics, University of Bonn, Bonn, Germany. ²⁴Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany. ²⁵Department of Neurology, Technische Universität München, Munich, Germany. ²⁶Institute of Human Genetics, Technische Universität München, Munich, Germany. ²⁷Institute of Human Genetics, Helmholtz Zentrum Munich, German Research Center for Environmental Health, Munich, Germany. ²⁸Department of Genetics, University Medical Center and University of Groningen, Groningen, The Netherlands. ²⁹Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. ³⁰Wellcome Trust Centre for Human Genetics, Oxford, UK. ³¹Department of Dermatology, University Hospital, Schleswig-Holstein, Christian-Albrechts-University, Kiel, Germany. ³²PopGen Biobank, University Hospital Schleswig-Holstein, Kiel, Germany. ³³Department of Physiology, Centre of Translational Medicine and Centre for Translational Genomics, University of Tartu, Tartu, Estonia. ³⁴Department of Dermatology and Venerology, University of Tartu, Tartu, Estonia. ³⁵Estonian Genome Center, University of Tartu, Tartu, Estonia. ³⁶Department of Dermatology, Henry Ford Hospital, Detroit, Michigan, USA. ³⁷Institute of Epidemiology I, Helmholtz Zentrum Munich, German Research Center for Environmental Health, Neuherberg, Germany. ³⁸Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-University, Munich, Germany. ³⁹Klinikum Grosshadern, Munich, Germany. ⁴⁰Department of Medicine, Division of Rheumatology, University of Toronto, Toronto Western Hospital, Toronto, Ontario, Canada. ⁴¹Department of Medicine, Division of Dermatology, University of Toronto, Toronto Western Hospital, Toronto, Ontario, Canada. ⁴²Department of Medicine, Memorial University, St. John's, Newfoundland, Canada. ⁴³Dermatological Sciences, Salford Royal NHS Foundation Trust, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. ⁴⁴Institute of Human Genetics, University of Erlangen-Nuremberg, Erlangen, Germany. ⁴⁵Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden. ⁴⁶Folkhälsan Institute of Genetics, Helsinki, Finland. ⁴⁷Department of Medical Genetics, University of Helsinki, Helsinki, Finland. ⁴⁸A list of members and affiliations appears at the end of the paper. ⁴⁹Ann Arbor Veterans Affairs Hospital, Ann Arbor, Michigan, USA. ⁵⁰Queen Mary University of London, Barts and the London School of Medicine and Dentistry, London, UK. ⁵¹These authors contributed equally to this work. ⁵²National Center for Biotechnology Information, National Library of Medicine, US National Institutes of Health, Bethesda, Maryland, USA. ⁵³Cardiovascular Research Institute, University of California, San Francisco, San Francisco, California, USA. ⁵⁴Center for Human Genetics, University of California, San Francisco, San Francisco, California, USA. ⁵⁵Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA. ⁵⁶Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, USA. ⁵⁷Celera, Alameda, California, USA. ⁵⁸Department of Pediatrics, University of Utah, Salt Lake City, Utah, USA. ⁵⁹Department of Rheumatology, University of Utah, Salt Lake City, Utah, USA. ⁶⁰Department of Gastroenterology, University of Utah, Salt Lake City, Utah, USA. ⁶¹Centre National de Génotypage, Institut Génomique, Commissariat à l'Energie Atomique, Evry, France. ⁶²Department of Dermatology, University of California, San Francisco, San Francisco, California, USA. ⁶³Department of Dermatology, Baylor University Medical Center, Dallas, Texas, USA. ⁶⁴Seay Center for Musculoskeletal Research, Texas Scottish Rite Hospital for Children, Dallas, Texas, USA. ⁶⁵Department of Dermatology and Venerology, University of Helsinki, Helsinki, Finland. ⁶⁶Asklepios Nordseeklinik, Westerland/Sylt, Germany. ⁶⁷TOMESA Clinics, Bad Salzschlirf, Germany. ⁶⁸Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. ⁶⁹Department of Dermatology, University of Göttingen, Göttingen, Germany. ⁷⁰Department of Dermatology, University of Münster, Münster, Germany. ⁷¹Psoriasis Rehabilitation Hospital, Bad Bentheim, Germany. ⁷²Department of Endocrinology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. ⁷³Department of Epidemiology and Biostatistics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. ⁷⁴University of Dundee, Dundee, UK. ⁷⁵Department of Dermatology, University College Hospital Galway, Galway, Ireland. ⁷⁶Department of Dermatology, St. Vincent's University Hospital, Dublin, Ireland. ⁷⁷Department of Clinical Medicine, Trinity College Dublin, Our Lady's Children's Hospital Crumlin, Dublin, Ireland. ⁷⁸Dermatology Service, Hospital del Mar-Instituto Municipal de Asistencia Sanitaria (IMAS), Barcelona, Spain. ⁷⁹Department of Dermatology and Venerology, Sahlgrenska University Hospital, Gothenburg, Sweden. ⁸⁰National Psoriasis Foundation, Portland, Oregon, USA. ⁸¹Institute of Genetic Epidemiology, Helmholtz Zentrum Munich, German Research Center for Environmental Health, Neuherberg, Germany. ⁸²Research Unit Molecular Epidemiology, Helmholtz Zentrum Munich, German Research Center for Environmental Health, Neuherberg, Germany. ⁸³Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University of Duisburg-Essen, Essen, Germany. ⁸⁴Clinic of Cardiology, West German Heart Centre, University Hospital of Essen, University Duisburg-Essen, Essen, Germany. ⁸⁵Department of Statistics, University of Oxford, Oxford, UK. ⁸⁶Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, Suciaco, Western Australia, Australia. ⁸⁷Genetics and Infection Laboratory, Cambridge Institute of Medical Research, Addenbrooke's Hospital, Cambridge, UK. ⁸⁸Division of Psychological Medicine and Psychiatry, Biomedical Research Centre for Mental Health at the Institute of Psychiatry, King's College London, London, UK. ⁸⁹The South London and Maudsley NHS Foundation Trust, Denmark Hill, London, UK. ⁹⁰Diamantina Institute of Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Brisbane, Queensland, Australia. ⁹¹Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine,

London, UK. ⁹²Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland. ⁹³Department of Psychological Medicine, Cardiff University School of Medicine, Heath Park, Cardiff, UK. ⁹⁴Molecular and Physiological Sciences, The Wellcome Trust, London, UK. ⁹⁵Centre for Gastroenterology, Barts and the London School of Medicine and Dentistry, London, UK. ⁹⁶Clinical Neurosciences, St. George's University of London, London, UK. ⁹⁷Oxford Centre for Diabetes, Endocrinology and Metabolism (ICDEM), Churchill Hospital, Oxford, UK. ⁹⁸Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee, UK. ⁹⁹Social, Genetic and Developmental Psychiatry Centre, King's College London Institute of Psychiatry, Denmark Hill, London, UK. ¹⁰⁰University of Cambridge Department of Clinical Neurosciences, Addenbrooke's Hospital, Cambridge, UK. ¹⁰¹Department of Cardiovascular Science, University of Leicester, Glenfield Hospital, Leicester, UK. ¹⁰²Glaucoma Research Unit, Moorfields Eye Hospital NHS Foundation Trust, London, UK. ¹⁰³Department of Genetics, University College London Institute of Ophthalmology, London, UK. ¹⁰⁴Department of Molecular Neuroscience, Institute of Neurology, London, UK. Correspondence should be addressed to R.C.T. (vp-health@qmul.ac.uk), G.R.A. (goncalo@umich.edu) or J.T.E. (jelder@umich.edu).

ONLINE METHODS

Sample collections. The samples used in the three GWAS data sets (Kiel, CASP and WTCCC2) were previously described^{3–5}. Samples in the PAGE and GAPC data sets (**Supplementary Tables 1 and 2**) were collected from subjects of European descent at the participating institutions after obtaining informed consent. Enrollment of human subjects for this study was approved by the ethics boards of the participating institutions in adherence with the Declaration of Helsinki Principles. DNA was isolated from blood or Epstein-Barr virus (EBV)-immortalized lymphoblastoid cell lines using standard methods.

The collections used in the GAPC and PAGE Immunochip studies are described in **Supplementary Table 2**.

The samples from GAPC substantially overlapped with those described as replication data sets in Strange *et al.*⁵. All cases had been diagnosed as having psoriasis vulgaris. The GAPC cases and the Irish and Spanish controls were genotyped at the Wellcome Trust Sanger Institute, and all samples were provided by the relevant groups listed in **Supplementary Table 2** and by the members of the GAP Consortium listed in the **Supplementary Note**. The UK controls were the WTCCC common control samples that did not overlap with the samples included in the original GWAS (the data set consisted of 6,740 1958 British Birth cohort and 2,900 UK Blood Service samples genotyped at the Wellcome Trust Sanger Institute and the University of Virginia). The German controls were obtained from the PopGen Biobank and genotyped at the Institute of Clinical Molecular Biology at Christian-Albrechts-University of Kiel. The Finland control data were from the DILGOM collection⁷⁰. The Irish controls were provided by the Irish Blood Transfusion Service/Trinity College Dublin Biobank, and Irish cases were collected with the aid of the Dublin Centre for Clinical Research. We did not include specific controls from Austria or Sweden, but principal-components analysis suggested that the cases from these cohorts were well matched to the controls from The Netherlands and Germany.

For the PAGE Immunochip study, samples also substantially overlapped with previously published replication data sets. The German cases (described as a replication data set in Ellinghaus *et al.*³) and all samples from the United States and Canada, as well as 439 Estonian cases from the University of Tartu, were genotyped at the Institute of Clinical Molecular Biology at Christian-Albrechts-University of Kiel. The respective samples were provided by the groups listed in **Supplementary Table 2** and by the members of PAGE listed in the **Supplementary Note**. The German controls were obtained from a population-based sample from the general population living in the region of Augsburg in southern Germany (KORA S4/F4)⁷¹, which was genotyped at the Helmholtz Zentrum Munich, and from the population-based epidemiological HNR study, for which genotyping was performed at the Life and Brain Center at the University Clinic in Bonn. The remaining Estonian samples were obtained from and genotyped at the Estonian Genome Center University of Tartu (EGCUT).

Genotyping panel and SNPs. The Immunochip is a custom Illumina Infinium high-density array consisting of 196,524 variants (after Illumina quality control) compiled largely from variants identified in previous GWAS of 12 different immune-mediated inflammatory diseases, including psoriasis¹³. The main aims of the Immunochip were deeper replication and fine mapping of genome-wide significant loci, as well as increasing power to promote promising but less significant SNPs to genome-wide significance. For fine mapping, SNPs within 0.2 cM on either side of the GWAS top SNPs at 186 loci were selected from 1000 Genomes Project⁷² low-coverage pilot Utah residents of Northern and Western European ancestry (CEU) sequencing data as well as additional variants identified by resequencing from groups involved in the chip design. For promotion of promising signals and those not quite reaching genome-wide significance, each disease-focused group was allowed to submit approximately 3,000 additional SNPs. We submitted 17 of the 19 confirmed regions associated with psoriasis at genome-wide significance (**Table 1**) for fine mapping on the basis of a preliminary meta-analysis of our data, and 1 of the confirmed signals (*IL28RA*) and 9 of the new psoriasis-associated signals (**Supplementary Table 5**, disease overlap) were submitted for fine-mapping by groups studying other diseases (although we also submitted 3 of these as part of our additional SNP allocation: rs11121129, rs10865331 and rs9504361). Six additional signals were detected on the basis of additional

SNP allocation in individual groups; four of these (rs11795343, rs4561177, rs11652075 and rs545979) were submitted by our group. All Immunochip samples were genotyped as described in Illumina's protocols.

Genotype calling. For the PAGE data set, genotype calling was performed using Illumina's GenomeStudio Data Analysis software and the custom-generated cluster file of Trynka *et al.*¹³. (which was generated by initial clustering of 2,000 UK samples with the GenTrain2.0 algorithm and subsequent manual readjustment and quality control). Genotype calling for the GAPC data set was performed using GenoSNP⁷³ from allele intensities, except for the German, Italian, Dutch and Finnish controls, which were called using the same method described for the PAGE data set.

Imputation. To increase the number of overlapping SNPs between data sets, we performed imputation on the 3 GWAS data sets using minimac⁷⁴ (Kiel and CASP) and IMPUTE2 (refs. 75,76) (WTCCC2) using data from CEU reference haplotypes from the 1000 Genomes Project⁷² (December 2010 version of the 10/08/04 sequence and alignment release containing 629 individuals of European descent). SNPs with low imputation quality ($r^2 \leq 0.3$ for minimac and info score < 0.5 for IMPUTE2) were removed. For all 3 data sets, cases and controls were imputed together.

Sample and genotype quality control. For the Immunochip data sets, we first excluded SNPs with a call rate below 95% or with a Hardy-Weinberg equilibrium P value of $< 1 \times 10^{-6}$. Samples with less than SNP call rates below 98% were then excluded. Because the Immunochip includes a large proportion of fine-mapping SNPs that are associated with autoimmune disease, we used a set of independent SNPs that have P values > 0.5 from the meta-analysis of the three GWAS data sets as a quality control tool for each individual Immunochip data set. Using the HapMap 3 samples as reference⁷⁷, we performed principal-component analysis to identify and remove samples with non-European ancestry. We also removed samples with extreme inbreeding coefficients or heterozygosity values computed by PLINK⁷⁸.

To assess possible stratification in the data sets, principal-components analysis was also performed in each of the Immunochip data sets separately (excluding HapMap). There was no evidence of stratification between the cases and controls of each sample group. However, as expected, the top principal components did separate the samples well by country of origin. The use of the top ten eigenvectors as covariates in the analysis did not completely correct for stratification, and, therefore, a linear mixed-model method (efficient mixed-model association expedited (EMMAX)) was instead used for the association analysis. These methods have been shown to outperform principal components in correcting for this type of population stratification and cryptic relatedness⁷⁹, which is becoming more common as sample sizes increase and studies comprise more collaborative efforts.

To identify duplicate pairs or highly related individuals among data sets, we used a panel of 873 independent SNPs that were genotyped in both the GWAS and Immunochip samples and performed pairwise comparisons using the genome function in PLINK⁷⁸, requiring $Pi-HAT$ of ≥ 0.5 . We identified 1,142 (885 from GAPC and 257 from PAGE) related sample pairs (mostly duplicates) and removed one sample from each pair. We also removed 4,828 controls from the UK common Immunochip controls because of duplication in the WTCCC2 GWAS sample. For GWAS samples that were duplicated in the Immunochip data sets (the majority of duplicates), we removed the samples from the Immunochip data sets to keep the previously published data sets intact.

The GWAS data sets underwent quality control as previously described and were analyzed for association using the top principal components from the previous analyses as covariates^{3–5}.

We visually inspected the signal intensity cluster plots for all SNPs with associations reaching genome-wide significance to confirm high-quality genotype calling.

Genomic control. Genomic control inflation factors for the five data sets were 1.09 (Kiel), 1.06 (CASP), 1.04 (WTCCC2), 0.99 (PAGE) and 0.96 (GAPC), indicating that population structure and cryptic relatedness were adequately controlled for in these data sets. Because the Immunochip was designed for deep replication and fine mapping of loci associated with autoimmune

diseases¹², using all independent SNPs from the chip would not give an accurate estimate of the genomic control⁸⁰ value (λ_{GC}). Therefore, we selected common SNPs (with minor allele frequency (MAF) of >0.05) from the ImmunoChip that had association *P* values of >0.5 on the basis of a meta-analysis combining the Kiel, CASP and WTCCC2 GWAS, and then performed LD pruning to identify an independent SNP set to compute λ_{GC} for the association results from the ImmunoChip data sets. As a result of SNP selection bias, the genomic control value for the final meta-analysis was computed using a set of independent SNPs associated with reading and writing ability (J.C. Barrett, personal communication). We further removed SNPs that were within 500 kb of previously detected psoriasis-associated loci (within 3 Mb in the MHC region), and the remaining 1,426 SNPs yielded λ_{GC} of 1.11 for the meta-analysis overall. Using λ_{1000} (ref. 81), the genomic control inflation factor for an equivalent study of 1,000 cases and 1,000 controls, the rescaled λ equaled 1.01.

70. Inouye, M. *et al.* An immune response network associated with blood lipid levels. *PLoS Genet.* **6**, pii e1001113 (2010).
71. Wichmann, H.E., Gieger, C. & Illig, T. KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* **67** suppl 1, S26–S30 (2005).

72. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
73. Giannoulatos, E., Yau, C., Colella, S., Ragoussis, J. & Holmes, C.C. GenoSNP: a variational Bayes within-sample SNP genotyping algorithm that does not require a reference population. *Bioinformatics* **24**, 2209–2214 (2008).
74. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).
75. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
76. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
77. Altshuler, D.M. *et al.* Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).
78. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
79. Kang, H.M. *et al.* Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42**, 348–354 (2010).
80. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
81. Freedman, M.L. *et al.* Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388–393 (2004).

Psoriasis: from bed to bench and back

Advances in immunology, a focus on human tissue and a dash of luck have given dermatologists several highly effective psoriasis therapies, of which Stelara is the latest. Ken Garber reports.

Psoriasis is a unique biotech success story. Four targeted biological therapies have received US Food and Drug Administration (FDA) approval since 2004. And even though the most recently approved product, Johnson & Johnson's Stelara (ustekinumab), a monoclonal antibody (mAb) against the p40 subunit of interleukin (IL)-12/23 (ref. 1), has exceptional effectiveness, psoriasis can still only be managed rather than cured. That's important both medically—the disease affects about 2% of people of European descent—and financially. The current worldwide market for psoriasis medications is \$3.3 billion, and is expected to steadily grow. Meanwhile, a wave of even more highly targeted therapies, both biologics and small molecules, is working its way through the development pipeline.

But it's not the effectiveness of the therapies that make the cache of psoriasis drugs unique; biologics have transformed the treatment of several chronic diseases. What makes these drugs special is that they have taken a unique trajectory to the clinic. Instead of following the canonical bench-to-bedside translational model—in which genetics indicate disease mechanisms, which are worked out and tested in *in vitro* systems and animal models, before drugs are brought into the clinic—researchers relied on a bed-to-bench-and-back approach. Lacking an accepted animal model for psoriasis, researchers resorted to deconvoluting the disease from human tissue specimens using the tools of cell biology and functional genomics, with a bit of luck added in the form of therapies that worked in unexpected ways. And the considerable insights human genetics has delivered have mostly confirmed what was already known or suspected.

"What happened is the combination of cell biology, expression profiling, genetics and therapeutics have really all converged," says psoriasis researcher James Krueger of Rockefeller University in New York. "The puzzle fits together very neatly." Krueger thinks this bedside-to-bench translational model could be applied to other autoimmune diseases and to cancer.

Right drug, wrong reason

T cells were identified as the main culprit in psoriasis about 20 years ago. Before that, the disease was thought to originate from

abnormalities in keratinocytes, the predominant cell type in the epidermis, the outer layer of the skin. Keratinocyte proliferation is a major feature of psoriasis, contributing to the scaly plaques that cover the skin of people with the disease, so it was a reasonable assumption. But the advent of immunohistochemistry in the 1980s allowed researchers to visualize T-cell infiltrates in psoriatic lesions from patients. And the effectiveness of immunosuppressants like cyclosporine and FK-506, which completely block the production of cytokines from T cells, tilted the consensus toward psoriasis as a T cell-mediated disease. "That... laid the ground for what we have now, which is the biologic era in psoriasis," says dermatology researcher Christopher Griffiths of the University of Manchester, UK.

That era has not been without its wrong turns. Biologics directly targeting T cells have been only modestly effective. Tumor necrosis factor blockers, such as the fusion protein Enbrel (etanercept; Amgen in Thousand Oaks, California) and the mAb Remicade (infliximab; Centocor of Horsham, Pennsylvania), the effectiveness of which in psoriasis became apparent only by accident, work much better, for reasons that are only now becoming clear. Another mistake was assuming that psoriasis was driven by T-helper 1 (T_H1) cells—an error that paradoxically led to the creation of the highly effective Stelara (Box 1).

The error originated when researchers found that T_H1 cells are abundant in psoriatic lesions. The cytokine IL-12, a heterodimer with two subunits, p35 and p40, drives T_H1 cell differentiation. "So we started to look for IL-12, and we found what we thought was IL-12 in psoriasis lesions, because the IL-12 p40 [subunit] mRNA

was very high," recalls Krueger. "So it looked like it was a T_H1 disease."

Centocor noticed, too, and decided to create an antibody against IL-12 to inhibit T_H1 cells. This was Stelara. What no one knew at the time was that IL-12's p40 subunit was shared with IL-23, a then-unknown cytokine essential for T_H17 function. It wasn't until Stelara entered the clinic that the company realized that the drug, which binds p40, also targets IL-23. IL-23 was discovered in 2000, but its role in T_H17 cell proliferation wasn't obvious until 2005, well into Stelara's clinical development. "It's the happy accident that this p40 subunit is shared with IL-23," says Krueger. "One of those really, really lucky happenings."

All the evidence now points to IL-23 and T_H17 cells driving the disease, not IL-12 and T_H1 cells. (T_H1 cells may play a secondary role.) Krueger's group in 2004 reported abundant IL-23 in psoriasis lesions², and four years later found a large number of cells there secreting IL-17 (ref. 3). Genetic validation followed. In 2007, human geneticists reported an association between variants of the IL-23 receptor gene and psoriasis⁴, and two years later they implicated IL-23A, the gene that encodes the IL-23 p19 subunit⁵. Finally, last year researchers at King's College London reported that an anti-IL-23 mAb blocks psoriasis in skin-transplanted mice⁶. "That's the best evidence so far that IL-23 is the real deal," says psoriasis researcher James Elder of the University of Michigan in Ann Arbor.

Thus, genetics and animal experiments only confirmed what researchers were guessing—that Stelara works in psoriasis because it attacks IL-23 and T_H17 cells, far from the company's original rationale for the drug. T_H17 cells are at the center of the current model of psoriasis pathogenesis. In simplified form, the model goes: dendritic cells in the skin present some unknown antigen, while prompting other dendritic cells to secrete cytokines, including IL-23, leading to the differentiation of T_H17 cells. These in turn produce cytokines like IL-17 and IL-22 that activate keratinocytes. The keratinocytes (and psoriatic lesions) proliferate and

Box 1 CD4⁺ cell lineage expansion

Stelara has proven spectacularly successful in psoriasis, even though the drug was developed against the wrong CD4⁺ cell lineage. Currently, CD4⁺ T cells are divided into three accepted lineages—T_H1, T_H2 and T_H17—based on the cytokines they produce and their distinct roles in host defense against pathogens. T_H1 and T_H2 cells were the first to be distinguished, with T_H1 cells mainly involved in facilitating cellular immunity, T_H2 humoral. Over the past five years, T_H17 cells have emerged as the major inflammatory mediator in autoimmune diseases. Two companies recently reported, to little fanfare, dramatic early results in clinical trials with psoriasis and an unexpected mechanism for mAbs targeting IL-17, the signature T_H17 cytokine, and its receptor.

produce a variety of cytokines including tumor necrosis factor alpha (TNF- α), chemokines and antimicrobial peptides that feed back into the inflammatory cascade, forming an autoinflammatory loop (Fig. 1). This model derives from immunohistochemical and microarray analyses of human psoriasis lesions, from therapeutics, and most recently from genome-wide association studies, which have implicated T_H17 cytokines as well as nuclear factor (NF)- κ B signaling downstream of TNF- α . “It’s remarkable how these [associated genetic] regions all align with the pathways that have been implicated by treatment,” says Elder.

Cascading antibody effects

Companies responded by quickly developing mAbs that intentionally target T_H17 cytokines specifically, not only for psoriasis but for rheumatoid arthritis, inflammatory bowel disease and other autoimmune diseases identified as T_H17 diseases. (These include multiple sclerosis, despite Stelara’s complete and unexpected failure to work in that disease.) This was done despite the fact that targeting a single cytokine in a complex disease carries a large risk of failure; indeed, many clinical trials are now underway (Table 1).

In psoriasis, two antibodies against the p19 subunit of IL-23 are in phase 2. Unlike Stelara, these specifically target T_H17 cells and thus are theoretically safer because they don’t shut down T_H1 immune responses to pathogens. But clinical results have yet to be reported.

Novartis, in Basel, and Amgen have reported early clinical data for their antibodies targeting IL-17A and its receptor, respectively. The strategy here was to target the cytokine and its receptor, rather than the T_H17 cell itself. Both drugs work. In the 18 psoriasis patients receiving a single infusion of Novartis’s AIN457 antibody, there was a 63% reduction in mean psoriasis area and severity index (PASI) scores and ~40% achieved a PASI75 score⁷. (PASI is a scale of disease severity based on extent of psoriatic coverage and severity of lesions and PASI75 means at least a 75% PASI improvement from baseline—the current standard for an excellent response.) These results were impressive, especially for an early trial, but not spectacular; in contrast, about two-thirds of Stelara patients in phase 3 achieved a PASI75. Amgen’s molecule, AMG 827, even though patient numbers so far are miniscule, surpassed this: seven of eight patients receiving a single dose achieved a PASI75.

Why are the Novartis and Amgen drugs so effective? Microarray studies of patient biopsy samples show a surprising cascade of drug effects. Novartis saw a reduction not only in IL-17A after treatment, as expected, but also

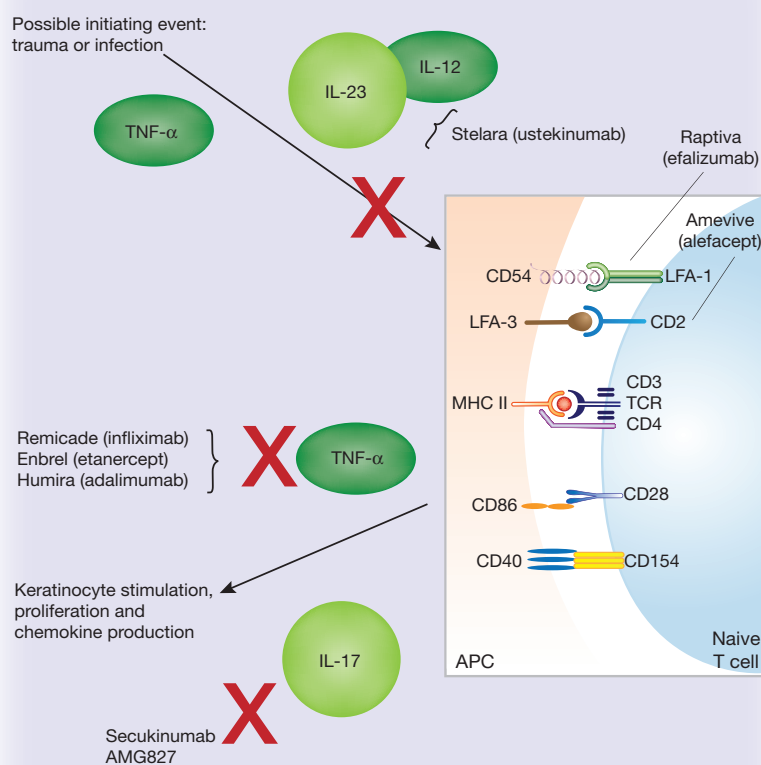


Figure 1 Wealth of targets. Stelara targets differentiation cytokines for T_H1 and T_H17, thus blunting inflammation and preventing abnormal keratinocyte differentiation and proliferation. Raptiva (withdrawn from the market in 2009) and Amevive (alefacept) target adhesion molecules at the immunological synapse, preventing T-cell activation. TNF blockers interfere with NF- κ B-mediated inflammation. And antibodies against IL-17 and its receptor interrupt a proinflammatory feedback loop. Not shown: JAK inhibitors interrupt JAK-Stat signaling downstream of many cytokine receptors, leading to incomplete but broad downregulation of innate and adaptive immunity. Source: Modified from FDA Dermatological and Ophthalmic Drugs Advisory Committee meeting, January 17, 2008.

surprisingly in transcripts for IL-12B, which codes for the p40 subunit of IL-23, an upstream cytokine. Just as unexpected were drops in likely IL-17-producing T cells, IL-22 and TNF- α . The Novartis mAb, although it targets only existing IL-17A cytokine, “affects the inflammatory response in a way that prevents additional IL-17A expression,” says Dhavalkumar Patel, head of autoimmunity, transplantation and inflammation at Novartis. Patel postulates that AIN457 is interrupting an inflammatory loop. “Inhibiting the product of a certain cell type, one wouldn’t expect the cell type, in this case T_H17, to go down, unless there were other factors and cascades involved,” Patel says. “And that’s what we showed in the study...it really affects inflammation as a whole.”

Amgen reached a similar conclusion based on its own microarray results⁸. “We see the skin inflammation in psoriasis as a branched,

positive-feedback loop rather than a one-directional cytokine cascade,” writes Chris Russell, medical director for molecular sciences at Amgen, in an e-mail. “If it can be blocked at a key point, then all the other components of the loop can also be affected,” he adds. AMG 827 treatment dramatically reduced expression not only of downstream effectors but of thousands of overexpressed genes, Russell notes, including key cytokines like IL-23 and TNF- α , and within two weeks T cells and dendritic cells began disappearing and lesions began thinning.

“I don’t think that anybody who looked at the animal models out there would have guessed that blocking IL-17 would have the kind of [broad] effects it seems to have,” says Krueger. If the inflammatory loop hypothesis proves true, and the clinical results hold up in larger trials, the implications are enormous. “It’s a very exciting thing scientifically,” says Krueger. “It creates a new therapeutic path in this disease, and it’s

Table 1 Selected agents in development for psoriasis

Company	Agent	Target	Stage
Centocor	Stelara (ustekinumab)	IL-12 and IL-23 shared p40 subunit	Registration
Pfizer	Tofacitinib (CP-690550)	JAK kinases	Phase 3
Novartis	AIN457 (secukinumab)	IL-17A	Phase 2
Lilly	LY2439821 mAb	IL-17A	Phase 2
Amgen	AMG 827 mAb	IL-17 receptor	Phase 2
Merck	SCH 900222 mAb	p19 subunit of IL-23	Phase 2
Centocor	CNTO 1959	p19 subunit of IL-23	Phase 2
Incyte	Ruxolitinib (topical)	JAK1 and JAK2 kinases	Phase 2
Pfizer	Fezakinumab (ILV-094)	IL-22	Phase 2
Novartis	AEB071	PKC	Phase 2

PKC, protein kinase C.

probably one that is going to be...at least theoretically safer than blocking T_H1 and T_H17 cells with something like ustekinumab, because it will be narrower in terms of what's being blocked."

Surprises, good and bad

Despite these theoretical advantages, anti-IL-17 agents are a long way from approval. And Stelara and TNF- α blockers have set a high bar for efficacy. It's puzzling that both should work so well in psoriasis—what Krueger calls the "80-80 paradox"—because each drug can turn off the disease in 80% of patients, at least in some studies. TNF- α is thought to be acting mainly on innate immunity, and Stelara's targets, IL-12 and IL-23, are master regulators of adaptive immunity, so it's counterintuitive that they should both be so effective. TNF- α , in fact, was only tried in psoriasis by chance when, in 2000, a woman was treated with Remicade for inflammatory bowel disease. Coincidentally, she also had psoriasis, which largely disappeared after a single infusion of the drug. Psoriasis clinical trials followed, and in 2006 the FDA approved Remicade for that disease.

The 80-80 paradox may have now been solved. Krueger's group recently has reported that IL-17, compared with TNF- α , has a modest effect on gene transcription when added to cultured keratinocytes. Even so, "when the keratinocyte sees both of these cytokines together, there is a huge outpouring of inflammatory gene products," Krueger says. So keratinocytes may need signaling down both pathways for full-blown psoriasis to develop⁹. Conversely, Krueger says, "suppressing either side might turn the inflammatory circuitry down to a low enough level that you no longer have clinically recognizable disease." Elder has new data showing similar synergistic effects of the two cytokines on keratinocytes, and speculates that TNF- α may cause proinflammatory gene transcription, whereas signaling downstream of the IL-17 receptor may stabilize the

transcribed mRNA. Krueger's group had earlier shown that TNF- α blockers strongly suppress the production of IL-23 and IL-17 by dendritic cells. None of this could have been anticipated when these drugs were first developed. "Isn't it amazing that IL-12/23 p40 and now the IL-17 drugs, on the one hand, and TNF- α drugs on the other hand, all work?" Elder says.

But there have been unpleasant surprises from psoriasis drugs, too. Lilly's anti-IL-23 antibody program was terminated in phase 2, due in part to "a nonclinical finding related to binding of the antibody to retinal tissue and potential difficulties in appropriate dosing beneath the skin," writes Lilly spokesperson Judy Moore in an e-mail. More concerning, because it implicates Stelara by association, is the fate of briakinumab, which—like Stelara—is a mAb targeting the p40 subunit of IL-12 and IL-23. Briakinumab, from Abbott Laboratories in Abbott Park, Illinois, was very effective in phase 3, achieving a PASI75 in >80% of patients in some trials. But, in one trial, seven major cardiovascular events were reported in the briakinumab group, compared with none in the placebo group. Regulatory authorities took note and, according to an Abbott US Securities and Exchange Commission filing, indicated "the need for further analysis and the potential for additional studies." On January 14 this year, Abbott withdrew its briakinumab applications from both the US and European regulatory agencies.

Stelara, despite a few cardiovascular events in phase 3, did not raise similar concerns, and presumably post-marketing data since its FDA approval in September, 2009, have also confirmed the drug's safety. In briakinumab's case, the side effects may have been particular to the molecular properties of the antibody, and not a class effect—no one knows. Griffiths has reviewed the safety data for Stelara¹⁰. "We could find no adverse signal there," he says. But clinicians are nevertheless prescribing Stelara with caution, in part because of the experience

with Genentech's Raptiva (efalizumab, a mAb against CD11), which was approved for psoriasis in 2003 but withdrawn in 2009 after three patients developed progressive multifocal encephalopathy, a lethal demyelinating disease of the brain. Despite Stelara's clean safety record, briakinumab "certainly has raised my sensitivity, and I will tell you I ask everybody who I start [on] this drug [Stelara] to take a baby aspirin a day to try to mitigate the potential [cardiac] downside," says Krueger.

Stelara has so far been mostly used in patients who have failed Enbrel or other TNF- α blockers, despite showing superior efficacy to Enbrel in the one head-to-head trial published so far¹¹. That's because the safety profile of TNF- α blockers is thoroughly understood after 13 years on the market and millions of users, whereas Stelara is barely out of the gate. This may change with time, especially as Stelara is given only once every three months, far less often than TNF- α blockers. Stelara "works fabulously—that's all I can tell you," says Krueger. "We have people whose lives we change, and they walk around and they look like they don't have psoriasis. This is true of some TNF antagonists also, that with a really remarkable response you have people walking around who have a totally different lifestyle and maybe a totally different employability profile than they did before their disease was treated."

JAK hits the road

The biggest challenge to both classes of biologics will come from small molecules, because the convenience of a pill, all else being equal, will probably trump any subcutaneous injection or intravenous infusion, regardless of dosing schedule. The p38 MAP kinase, a master regulator of inflammation that activates TNF- α , IL-1 and IL-8 simultaneously, was an attractive target until toxicity put an end to most clinical trials of p38 inhibitors in autoimmune disease. A small-molecule inhibitor of protein kinase C, which is crucial for IL-2 production, and thus T-cell propagation, and may be important for T-cell activation, is in phase 2 at Novartis.

But by far the most promising targets on the horizon for psoriasis—and for other autoimmune diseases—are the Janus or 'JAK' kinases. New York-based Pfizer's JAK inhibitor tofacitinib is furthest along. Phase 3 efficacy results in rheumatoid arthritis have been roughly comparable with those from TNF- α blockers Enbrel, Remicade and Humira¹². "We've got [tofacitinib] forecasted as a blockbuster product globally," says Seamus Fernandez, a pharmaceuticals analyst at Leerink Swan in Boston. "It certainly has megablockbuster potential."

Tofacitinib and other JAK inhibitors have surprised researchers both with their efficacy and

relative lack of toxicity. JAK kinases are crucial in immunity. There are about 200 known cytokines, and roughly 60 of them signal through the so-called 'type I' and 'type II' cytokine receptors, which have no catalytic kinase activity of their own. These receptors instead rely on JAK kinases for receptor phosphorylation, which creates a docking site for downstream signaling molecules, including Stat transcription factors.

There are four JAKs, with JAK3 associating with the common gamma (γ) chain, the third receptor in the IL-2 receptor complex. Pfizer designed tofacitinib as a JAK3 inhibitor. The gamma c chain is shared with five other cytokine receptors, none of them receptors for the signature T_H17 cytokines. (Tofacitinib was synthesized well before the T_H17 era.) "We thought [tofacitinib] would be largely blocking IL-2," says National Institutes of Health immunologist John O'Shea, who cloned JAK3. "We envisioned it in principle as being an oral dactilizumab [Zenapax]." Zenapax, Roche's anti-IL-2 receptor antibody, is used to prevent organ transplant rejection.

But tofacitinib does much more. In recent experiments in cultured T cells¹³, O'Shea found that tofacitinib also blocked production of T_H17 cytokines in a subset of T cells that had earlier been shown to be highly pathogenic in a mouse model of multiple sclerosis. The drug also impaired signaling from several key cytokines that do not signal through gamma c chain receptors. Finally, in a mouse model of sepsis, the drug countered TNF, IL-6 (a key inflammatory cytokine and validated drug target) as well as other innate immune responses. The reason for these unexpected effects, says O'Shea, is that tofacitinib blocks JAK1, and to some extent JAK2, in addition to JAK3. "That's probably

what accounts for the efficacy that's being seen in phase 2 and phase 3 trials," he says.

But that raises the question of why tofacitinib isn't more toxic. After all, JAK3 mutations cause a severe combined immune deficiency (SCID)-like syndrome, and JAK1 and JAK2 knockouts are perinatal lethal and embryonic lethal, respectively. But so far the main toxicity concern for tofacitinib has been a modest rise in cholesterol levels in individuals taking the drug. O'Shea speculates that the drug may not be completely knocking down its targets, or that its pharmacokinetics may create periods of low drug exposure. "You may not be suppressing cytokine signaling for 24 hours a day," he says.

Tofacitinib is in three separate phase 3 trials in psoriasis. To date, only phase 1 results have been reported¹⁴; patients receiving the second-highest dose of the drug had a mean drop in PLSS score of 72%, compared with 12% for placebo. "We have biologic-like efficacy, manageable safety and the convenience of an oral form," said Geno Germano, president and general manager of Pfizer's specialty care and oncology business units, at the JP Morgan Healthcare Conference in January. Leerink Swan's Fernandez tempers his own enthusiasm with concerns over long-term safety, including cholesterol elevation, patient acceptance of a twice-a-day pill over infrequent infusions or injections, and Medicare reimbursement policy decisions. But he calls the drug "potentially transformational."

So many good psoriasis drugs are in clinical trials or already on the market that it will take years to sort out their relative benefits. "I expect another six drugs to be put forward for approval in the next five years," says Krueger. Meanwhile, human genetic studies continue to be produc-

tive, identifying new susceptibility loci for the disease and potential new targets¹⁵. Functional studies of these variants are now underway that should further clarify the pathogenesis of the disease, although its cause remains elusive. So does a cure. "We may not be able to cure it," says Griffiths. "[But] if we can identify people who are at risk of developing psoriasis...we might be able to switch it off. So that's the hope."

Meanwhile, psoriasis offers an example to other disease areas of what can be accomplished through bedside-to-bench research. It may not become a trend, as model systems are entrenched. But an approach based on human tissue "will only help to accelerate the path of therapeutic discovery," Krueger says. "Hopefully, having another example or two of success will help to cement [its] sensibility."

Ken Garber, Ann Arbor, Michigan

1. Benson, J.M. *et al. Nat. Biotechnol.* **29**, 615–624 (2011).
2. Lee, E. *et al. J. Exp. Med.* **199**, 125–130 (2004).
3. Lowes, M.A. *et al. J. Invest. Dermatol.* **128**, 1207–1211 (2008).
4. Cargill, M. *et al. Am. J. Hum. Genet.* **80**, 273–290 (2007).
5. Nair, R.P. *et al. Nat. Genet.* **41**, 199–204 (2009).
6. Tonel, G. *et al. J. Immunol.* **185**, 5688–5691 (2010).
7. Hueber W. *et al. Sci. Transl. Med.* **2**, 52raa72 (2010).
8. Russell C.B. *et al., J. Invest. Dermatol.* **130**, S46, abstract 273 (2010).
9. Chiricozzi, A. *et al. J. Invest. Dermatol.* **131**, 677–687 (2011).
10. Reich, K. *et al. Br. J. Dermatol.* **164**, 862–872 (2011).
11. Griffiths, C.E. *et al. N. Engl. J. Med.* **362**, 118–128 (2010).
12. Garber, K. *Nat. Biotechnol.* **29**, 467–468 (2011).
13. Ghoreschi, K. *et al. J. Immunol.* **186**, 4234–4243 (2011).
14. Boy, M.G. *et al. J. Invest. Dermatol.* **129**, 2299–2302 (2009).
15. Ellinghaus, E. *et al. Nat. Genet.* **42**, 991–995 (2010).



Photodisc/Getty

IMMUNE REGULATION

Local government

Commensal bacteria have crucial roles in shaping host immune responses, but most studies have focused on the intestinal microbiota and not on the microbial communities associated with other barrier tissues. Yasmine Belkaid and colleagues now report that commensals found in the skin modulate local T cell responses in a site-specific manner and promote protective immunity to a cutaneous parasite.

In initial experiments, the authors assessed the cutaneous immune cell compartment in mice that were housed under specific pathogen free (SPF) or germ-free conditions. Compared with SPF mice, germ-free animals contained higher frequencies of FOXP3⁺ regulatory T cells in the skin. In addition, T cells isolated from the skin of germ-free mice produced lower levels of interferon- γ (IFN γ) and interleukin-17A (IL-17A) following *ex vivo* stimulation.

Previous studies have shown that colonization of the intestine by segmented filamentous bacteria (SFB) promotes the production of IFN γ and IL-17A by intestinal T cells. Although monocolonization of germ-free mice with SFB increased the levels of these cytokines in the

intestine, it did not promote their production in the skin. Furthermore, treatment of SPF mice with oral antibiotics altered their intestinal microbiota, but did not affect the composition of the skin microbiota or effector cytokine production by cutaneous T cells. By contrast, monocolonization of germ-free mice with the skin commensal *Staphylococcus epidermidis* increased IL-17A levels in the skin but not in the intestine. Thus, commensals that have specifically adapted to the skin or the intestine seem to have unique roles in shaping the T cell responses at each respective site.

To further explore the effects of the microbiota on cutaneous immune responses, the authors used a model of dermal infection with the parasite *Leishmania major*. Germ-free mice had impaired effector T cell responses following infection with *L. major*, but this defect could be reversed if they were monocolonized with *S. epidermidis* at the time of infection. Interestingly, these experiments also showed that the skin pathology associated with *L. major* infection is dependent on the presence of commensals and not directly induced by the parasite.

“commensals found in the skin modulate local T cell responses in a site-specific manner”

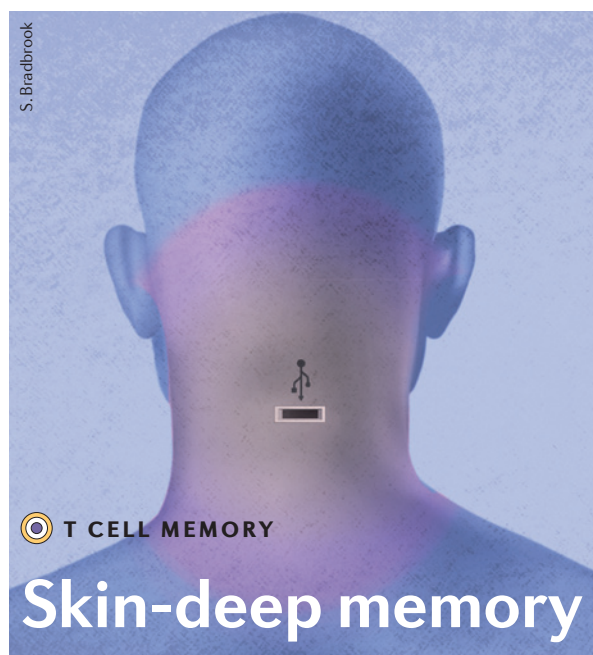


So how are skin commensals able to promote effector T cell responses? The authors found that these commensals can regulate various aspects of IL-1 signalling. Cells isolated from the skin of germ-free mice secreted lower levels of IL-1 α compared with cutaneous cells from SPF mice, but monocolonization of germ-free mice with *S. epidermidis* restored IL-1 α production to the levels seen in the SPF mice. In addition, keratinocytes from germ-free mice expressed higher levels of the mRNA encoding IL-1 receptor antagonist. Using gene-deficient mice, the authors showed that signalling via IL-1 receptor 1 (IL-1R1) and its downstream adaptor MYD88 drives IL-17A production by cutaneous T cells, but does not appear to be necessary for IL-17A production by intestinal T cells. Following culture with IL-1 α or IL-1 β , T cells purified from the skin and activated via their T cell receptors showed increased IL-17A production, suggesting that commensal-induced IL-1 can have direct effects on T cells. Finally, the authors found that in germ-free mice monocolonized with *S. epidermidis*, neutralization of IL-1 impaired the ability of these commensals to promote effector T cell responses during *L. major* infection.

Taken together, these findings suggest that commensals found in the skin, but not those in the intestine, augment IL-1 signalling to promote effector T cell responses locally. Notably, IL-1 signalling has been implicated in the pathology of psoriasis and other cutaneous disorders. Therefore, although skin commensals can promote protective immunity during infection, it is likely that the amplification of IL-1 signalling by local commensals may also exacerbate inflammatory skin diseases.

Yvonne Bordon

ORIGINAL RESEARCH PAPER Naik, S. et al. Compartmentalized control of skin immunity by resident commensals. *Science* 26 Jul 2012 (doi:10.1126/science.1225152)



Protective memory T cells have long been thought to reside in the blood and lymph nodes, but a new study in *Nature* shows that a population of memory CD8⁺ T cells exists permanently in the skin, where these cells provide protection against cutaneous re-infection. In a related paper published in *Science Translational Medicine*, it was shown that these skin-resident memory T cells, but not circulating malignant T cells, were spared in patients with T cell malignancies who were treated with the T cell-depleting antibody alemtuzumab and could continue to mediate immune protection in the skin.

“ skin-resident T_{EM} cells protect against pathogens even in the absence of T cell recruitment from the circulation ”

An accumulation of CD8⁺ T cells in the skin was observed following localized skin infection of mice with vaccinia virus. After resolution of cutaneous infection (day 30), virus-specific central memory CD8⁺ T (T_{CM}) cells were detected in the lymph nodes and effector memory CD8⁺ T (T_{EM}) cells were present in the skin. When these infected mice were surgically joined to mice that had not been previously infected to generate parabiotic pairs, the numbers of T_{CM} cells in the spleen and lymph nodes were similar in both parabionts, which indicates that these cells circulate in the blood. However, no T_{EM} cells were detected in the skin of the non-infected parabiont, even after prolonged periods, which indicates that T_{EM} cells are non-circulating, long-lasting residents of the skin.

To determine the relative roles of the two memory T cell populations in protection from infection, the parabionts were separated and then challenged with vaccinia virus infection of the skin. The virus was cleared much more efficiently in previously infected mice, which had both T_{CM} cells and skin-resident T_{EM} cells, than in the parabionts that had not been previously infected, which had only T_{CM} cells. Interestingly, skin-resident T_{EM} cells were also highly effective at rapidly eliminating the virus at distant skin sites, which indicates that they populate the entire skin surface and provide protection against re-infection independently of T_{CM} cells.

These findings have clinical implications for the treatment of leukaemic cutaneous T cell lymphoma (L-CTCL), which is a malignancy of T_{CM} cells, and of mycosis fungoides, which is a malignancy of skin-resident T_{EM} cells. Clark *et al.* report that low-dose alemtuzumab (a monoclonal antibody specific for CD52) effectively treats patients with L-CTCL but not mycosis fungoides. Indeed, alemtuzumab depleted all T cells from the blood, including skin-homing T_{CM} cells, but did not deplete skin-resident T_{EM} cells. Failure of alemtuzumab to deplete skin-resident T_{EM} cells was not due to poor skin penetration of the drug or loss of expression of CD52 by T_{EM} cells, but is thought to result from the scarcity of neutrophils in the skin to mediate T cell depletion through antibody-dependent cell-mediated cytotoxicity. Importantly, alemtuzumab-treated patients did not suffer from increased infections, which is consistent with the idea that skin-resident T_{EM} cells protect against pathogens even in the absence of T cell recruitment from the circulation.

Lucy Bird

ORIGINAL RESEARCH PAPERS Jiang, X. *et al.* Skin infection generates non-migratory memory CD8⁺ T_{RM} cells providing global skin immunity. *Nature* **483**, 227–231 (2012) | Clark, R. A. *et al.* Skin effector memory T cells do not recirculate and provide immune protection in alemtuzumab-treated CTCL patients. *Sci. Transl. Med.* **4**, 117ra7 (2012)

Skin infection generates non-migratory memory $CD8^+$ T_{RM} cells providing global skin immunity

Xiaodong Jiang¹, Rachael A. Clark¹, Luzheng Liu¹, Amy J. Wagers², Robert C. Fuhlbrigge¹ & Thomas S. Kupper¹

Protective T-cell memory has long been thought to reside in blood and lymph nodes, but recently the concept of immune memory in peripheral tissues mediated by resident memory T (T_{RM}) cells has been proposed^{1–5}. Here we show in mice that localized vaccinia virus (VACV) skin infection generates long-lived non-recirculating $CD8^+$ skin T_{RM} cells that reside within the entire skin. These skin T_{RM} cells are potent effector cells, and are superior to circulating central memory T (T_{CM}) cells at providing rapid long-term protection against cutaneous re-infection. We find that $CD8^+$ T cells are rapidly recruited to skin after acute VACV infection. $CD8^+$ T-cell recruitment to skin is independent of $CD4^+$ T cells and interferon- γ , but requires the expression of E- and P-selectin ligands by $CD8^+$ T cells. Using parabiotic mice, we further show that circulating $CD8^+$ T_{CM} and $CD8^+$ skin T_{RM} cells are both generated after skin infection; however, $CD8^+$ T_{CM} cells recirculate between blood and lymph nodes whereas T_{RM} cells remain in the skin. Cutaneous $CD8^+$ T_{RM} cells produce effector cytokines and persist for at least 6 months after infection. Mice with $CD8^+$ skin T_{RM} cells rapidly cleared a subsequent re-infection with VACV whereas mice with circulating T_{CM} but no skin T_{RM} cells showed greatly impaired viral clearance, indicating that T_{RM} cells provide superior protection. Finally, we show that T_{RM} cells generated as a result of localized

VACV skin infection reside not only in the site of infection, but also populate the entire skin surface and remain present for many months. Repeated re-infections lead to progressive accumulation of highly protective T_{RM} cells in non-involved skin. These findings have important implications for our understanding of protective immune memory at epithelial interfaces with the environment, and suggest novel strategies for vaccines that protect against tissue tropic organisms.

$CD8^+$ T cells have a pivotal role in antiviral immunity in target tissues^{6–9}. We infected the skin of control, $CD4^{-/-}$, or $CD4^+$ T-cell-depleted mice with VACV and assessed VACV-specific pentamer⁺ $CD8^+$ T cells¹⁰. Absence of $CD4^+$ T cells did not impair either antigen-specific $CD8^+$ T-cell proliferation in draining lymph nodes or subsequent accumulation in skin; in fact, the latter was enhanced (Fig. 1a, b). We then infected mice infused with OT-I ($CD8^+$) and OT-II ($CD4^+$) T cells with an ovalbumin-expressing VACV (VACV-Ova)¹¹. OT-I T cells are a transgenic $CD8^+$ T-cell population that recognize ovalbumin residues 257–264 in the context of H-2K^b, whereas OT-II T cells are a transgenic $CD4^+$ T-cell population recognizing ovalbumin residues 323–339 in the context of I-A^b (ref. 1). After skin infection, both OT-I and OT-II cells proliferated similarly in draining lymph nodes, and OT-I cells but not OT-II cells accumulated

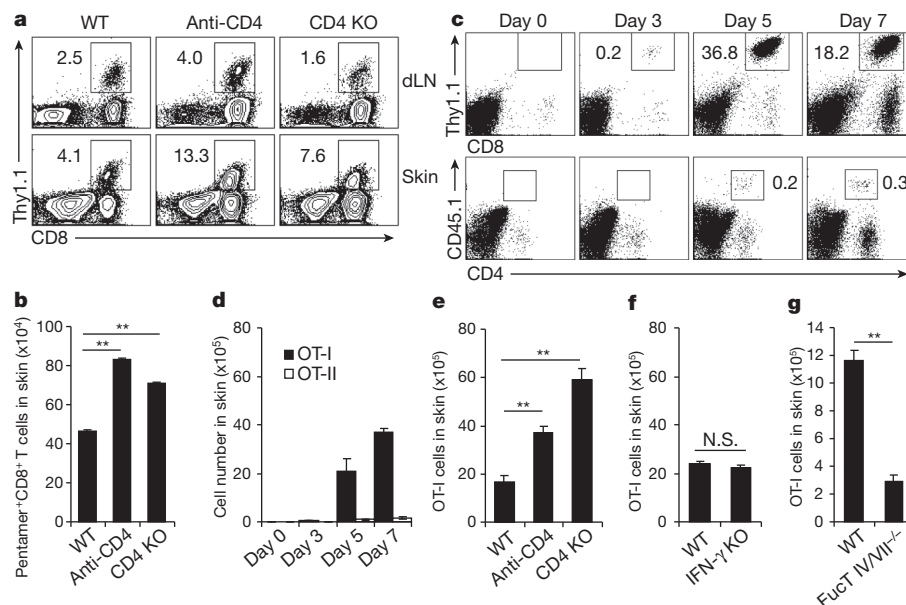


Figure 1 | $CD4^+$ T cells and IFN- γ are not required for acute recruitment of $CD8^+$ T cells to VACV-infected skin. **a, b**, Pentamer⁺ $CD8^+$ T cells in draining lymph nodes (dLN) and infected skin 7 days after VACV infection. KO, knockout; WT, wild type. **a**, The percentages of pentamer⁺ $CD8^+$ T cells are shown (excluding $CD19^+$ cells before flow cytometry analysis). **b**, The numbers of pentamer⁺ $CD8^+$ T cells in infected skin. **c, d**, The kinetic

infiltration of Thy1.1⁺ OT-I and CD45.1⁺ OT-II cells to infected skin after their co-transfer to naive mice. **e–g**, The numbers of OT-I cells in infected skin 7 days after infection in the absence of CD4, IFN- γ or Fc γ R IV/VII, respectively. All data are representative of at least three independent experiments ($n = 5$ mice per time point per group). **b, d–g**, Error bars show standard error of the mean (s.e.m.); ** $P < 0.01$; N.S., not significant.

¹Department of Dermatology and Harvard Skin Disease Research Center, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ²Department of Stem Cell and Regenerative Biology, Harvard University, Howard Hughes Medical Institute, Harvard Stem Cell Institute, Joslin Diabetes Center, Boston, Massachusetts 02115, USA.

significantly in infected skin (although other $CD4^+$ T cells showed some accumulation) (Supplementary Fig. 1 and Fig. 1c, d). Interestingly, OT-I cells accumulated in infected skin efficiently in the absence of either $CD4^+$ T cells or interferon (IFN)- γ (Fig. 1e, f), in contrast to a recently reported herpes simplex virus (HSV) vaginal infection model¹². However, skin accumulation (but not lymph-node proliferation) of OT-I cells from FucT IV/VII^{-/-} mice, which cannot make E- and P-selectin ligands, was significantly impaired (Fig. 1g and Supplementary Fig. 2a). Both E- and P-selectin were significantly upregulated in VACV-infected skin (Supplementary Fig. 2b). Thus, $CD8^+$ T-cell accumulation in skin after VACV infection does not require $CD4^+$ T cells or IFN- γ , but does require expression of E- and P-selectin ligands.

Murine models of viral infections of skin and other tissues have been useful in the study of T-cell memory^{13–16}. We explored the ability of $CD8^+$ memory T cells generated by VACV infection to recirculate after resolution of the cutaneous infection. We infected the skin of mice infused with OT-I cells with VACV, and waited until complete resolution of the infection (30 days). At 30 days we could identify T_{CM} cells in lymph nodes and effector memory T (T_{EM}) cells in skin (Supplementary Fig. 3a, b). We then surgically created parabiotic pairs between the infected mice and never-infected naive mice that had not been given OT-I cells. Parabiotic pairs were maintained for 2, 4, 8, 12 and 24 weeks, at which point they were surgically separated for the analysis of VACV-specific OT-I T cells (Fig. 2a). Mice joined for 8 weeks had similar numbers of OT-I T_{CM} cells in the spleen and lymph nodes of both parabionts, indicating rapid recirculation and

equilibration of T_{CM} cells (Fig. 2b, d). However, at 2–24 weeks there were no OT-I T_{RM} cells in the skin of the unimmunized parabiont (Fig. 2c, e). These early kinetics of T_{CM} recirculation and T_{RM} non-recirculation were confirmed by parabiotic mice that received no OT-I cells, using pentamer expression to identify VACV-specific memory cells (Supplementary Fig. 4). OT-I T_{RM} cells were readily identified in the skin of previously infected parabionts and mice that had been infected in parallel but never joined. These OT-I T_{RM} cells represented a significant fraction of total skin cells (Fig. 2c), and persisted for long periods of time: 30% of skin $CD8^+$ T cells at 12 weeks, and 15–20% of skin $CD8^+$ T cells at 24 weeks. In contrast, naive mice joined to previously infected mice had no OT-I T_{RM} cells in the skin, even after 24 weeks of parabiosis (Fig. 2c, e). Thus, skin T_{RM} cells persisted in skin for at least 6 months after infection and did not recirculate appreciably. T_{RM} and T_{CM} cells were further analysed by fluorescence-activated cells sorting (FACS) for expression of CD69, CD103, E- and P-selectin ligands, and production of IFN- γ and TNF- α upon activation. A subset of T_{RM} cells expressed BCL2, and T_{RM} cells lacked CD122 and CD127 (Supplementary Fig. 5). Immunofluorescence staining showed that many T_{RM} cells localize in epidermal and follicular epithelium, as reported in skin HSV infection¹⁷, but also localize in the dermis (Fig. 2f).

To compare directly the ability of T_{CM} and T_{RM} cells to eliminate a subsequent VACV infectious challenge, we infused μ MT mice with OT-I cells and infected them through the skin with VACV-Ova. After 30 days, mice were joined parabiotically to naive μ MT mice for 8 weeks, and then surgically separated. The infected parabiont contained

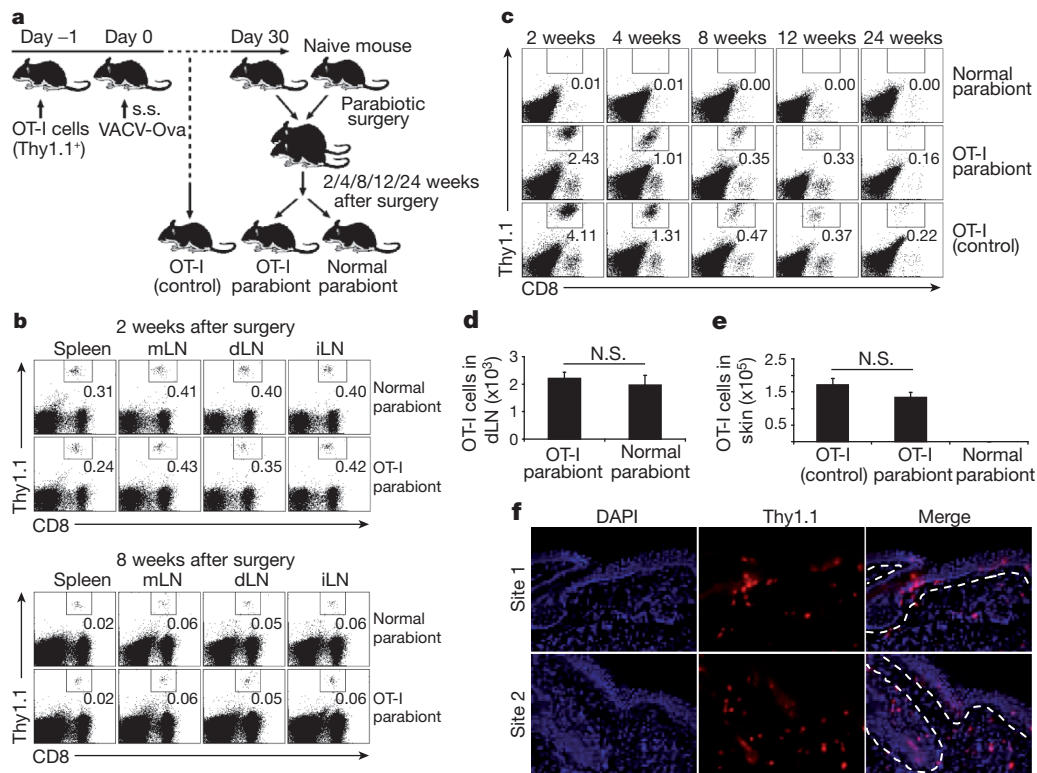


Figure 2 | $CD8^+$ T_{CM} cells recirculate quickly between parabiotic mice, but skin $CD8^+$ T_{RM} cells remain in place long term. **a**, 2×10^6 Thy1.1⁺ OT-I cells were intravenously transferred into Thy1.2⁺ mice 1 day before 2×10^6 plaque-forming units (p.f.u.) VACV-Ova skin scarification (s.s.). Thirty days after infection, OT-I-bearing mice were joined surgically with naive Thy1.2⁺ mice to create parabiotic mice. At indicated time points after surgery, parabiotic mice were separated and lymphoid tissues and skin were harvested to examine OT-I cells. **b**, The percentages of memory OT-I cells in the indicated lymphoid tissues between parabiotic mice 2 and 8 weeks after surgery were examined. mLN, mesenteric lymph node; dLN, draining lymph node; iLN, inguinal lymph

node. **c**, The percentages of OT-I T_{RM} cells in the skin of parabiotic and control mice over time were also examined. **d**, The numbers of OT-I T_{CM} cells in draining lymph nodes of parabiotic mice 8 weeks after surgery. **e**, The numbers of OT-I T_{RM} cells in the skin of parabiotic and control mice 8 weeks after surgery. **b–e**, Data are representative of three independent experiments ($n = 5$ mice per time point per group). **d**, **e**, Error bars, s.e.m.; N.S., not significant. **f**, Immunofluorescence of OT-I cells in infected skin sites 45 days after infection. Sections were stained for nuclei (4',6-diamidino-2-phenylindole (DAPI), blue) and Thy1.1 (red). $n = 15$ sections from 5 mice. Original magnification, $\times 200$. Two representative sites are shown.

T_{CM} cells in both the spleen and lymph nodes and T_{RM} cells within the skin, whereas the uninfected parabiont contained T_{CM} cells only in the spleen and lymph nodes (Fig. 2). At 2 weeks after separation, the skin of these mice was challenged with VACV-Ova, and assessed 6 days later for viral load (Fig. 3a). Despite the presence of abundant circulating OT-I T_{CM} cells, uninfected parabionts cleared the virus only 30-fold more effectively than naive mice (Fig. 3b). In contrast, the infected parabiont cleared the virus completely, 10^4 -fold more effectively than

the uninfected parabiont (Fig. 3b). Viral clearance was efficient even in mice treated with FTY720, a S1P inhibitor that blocks egress of T_{CM} cells from lymph nodes into blood¹ (Fig. 3b). To show that this was not an artefact of OT-I cell transfer, we reproduced and extended the experiment in a parabiotic model not involving transfer of OT-I cells (Fig. 3c). Parabiotic pairs were separated at 4 weeks, challenged with VACV, and assessed for viral load at 6, 14 and 26 days after challenge. There were again striking differences in the ability of endogenous T_{RM}

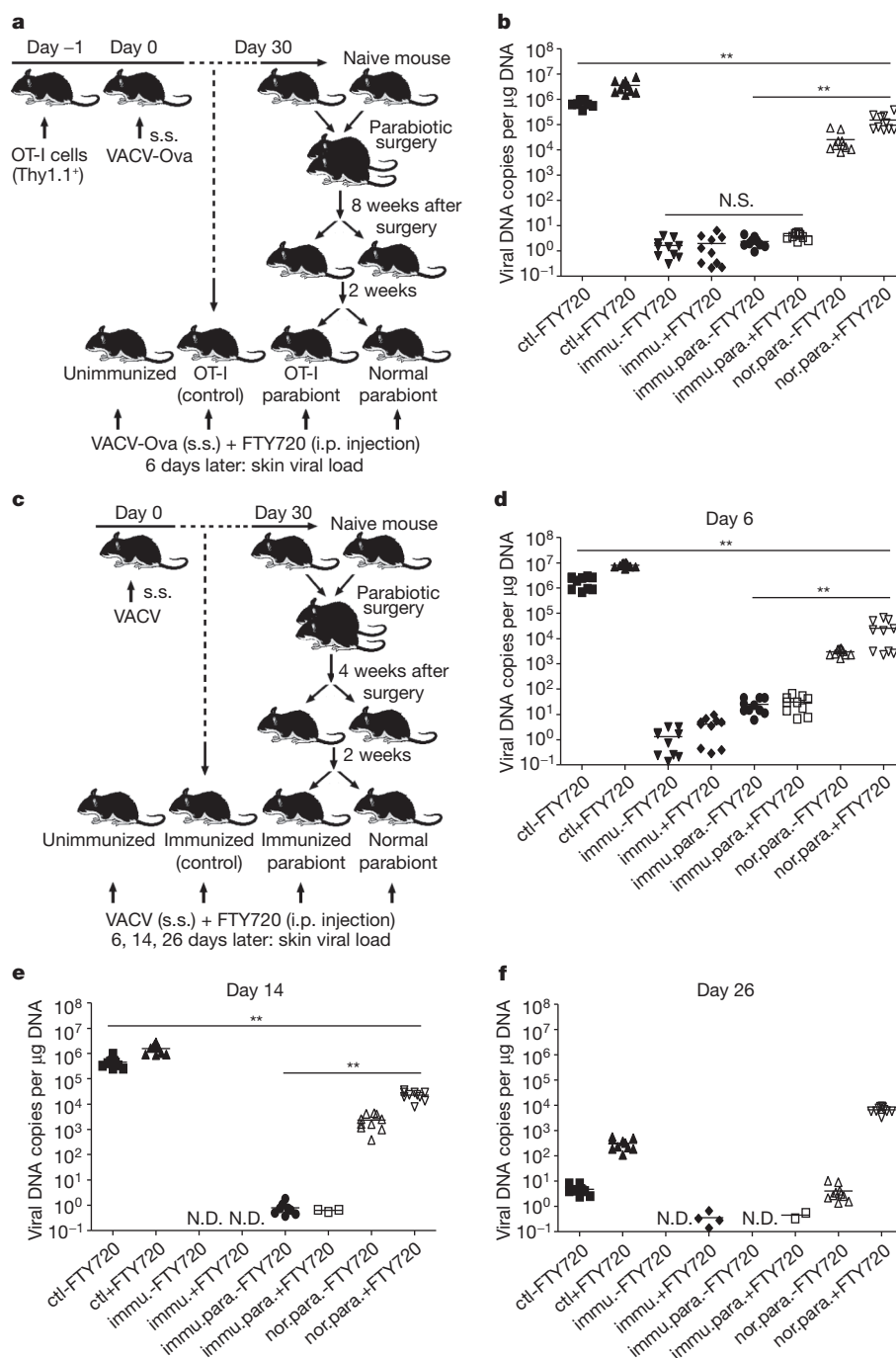


Figure 3 | Skin $CD8^+$ T_{RM} cells are superior to T_{CM} cells at protecting against re-infection. **a**, μ MT mice with OT-I transfer and VACV-Ova skin infection were used to create OT-I:normal parabiotic mice as described in Fig. 2. Eight weeks after surgery, parabiotic mice were separated. Two weeks later, separated mice were challenged with VACV-Ova on the skin. Half mice were injected daily with FTY720. **b**, Six days after challenge, skin viral load was assayed by quantitative polymerase chain reaction (qPCR). ctl, control; immu., immunized; nor., normal; para., parabiont. **c**, In separate experiments, μ MT

mice without OT-I transfer but with VACV skin infection were used to create immunized:unimmunized parabiotic mice. Four weeks after surgery, the same VACV skin challenge protocol was applied. **d–f**, Six, fourteen or twenty-six days after challenge, viral load was assayed. The results from duplicate qPCR runs are plotted. Horizontal bars indicate the mean. Data are representative of two independent experiments ($n = 5$ mice per group). ** $P < 0.01$; N.S., not significant; N.D., not detectable.

and T_{CM} cells to mediate viral clearance at days 6 (Fig. 3d) and 14 (Fig. 3e), which began to normalize by day 26 (Fig. 3f). However, FTY720-treated T_{CM} mice (the parabiont containing T_{CM} cells but no T_{RM} cells) are unable to clear virus even 26 days after infection, a time point by which naive mice have cleared the skin infection (Fig. 3f), presumably by generating protective T_{RM} cells. Thus, although T_{CM} cells are superior to naive T cells, they are inferior to T_{RM} cells at mediating rapid viral clearance from skin.

E- and P-selectin, CCL17 and ICAM1 are expressed constitutively on the blood vessels of normal skin^{18–20} and can support entry into uninfected skin of T_{EM} cells generated by cutaneous VACV infection. To study this phenomenon, we infected the left ears of OT-I-loaded mice with VACV-Ova and then measured the accumulation of OT-I cells in both infected (left) and uninfected (right) ears. There was measurable accumulation of OT-I cells in both ears, with similar kinetics. The absolute number of OT-I cells was always higher in the infected ear, but even 30 days after infection, OT-I cells represented a measurable fraction of all cells present in the uninfected ear (Fig. 4a). Thus, VACV skin infection generates $CD8^+ T_{RM}$ cells that distribute to distant skin sites as well as the site of infection. The accumulation of OT-I T_{RM} cells in distant skin sites is increased further after multiple sequential infections to other sites of skin (Fig. 4b–d), suggesting that skin T_{RM} cells continue to accumulate throughout skin in response to repeated cutaneous infections at distant sites.

To determine if $CD8^+ T_{RM}$ cells in distant skin sites were as protective as those at previously infected sites, we challenged μ MT mice previously infected on one ear with a second VACV skin infection on both ears either 7 or 30 days after the initial infection. FTY720 was administered to limit the contribution of T_{CM} cells (Fig. 4e). Notably, at both day 7 and day 30, T_{RM} cells in distant skin sites markedly reduced viral loads to levels comparable to those observed at the actual

site of previous infection, indicating that these distant T_{RM} cells were highly effective at rapidly eliminating virus. In contrast, viral loads in the skin of intraperitoneally (i.p.) immunized mice were between 10^3 and 10^4 higher at these time points (Fig. 4f, g). Thus, skin infection with VACV generates populations of T_{EM} cells that distribute to the entire skin surface, become T_{RM} cells, and mediate protection of the skin against re-infection with VACV in the absence of antibodies or T_{CM} cells.

T_{RM} cells have now been identified in the skin, gut, lung and brain in murine models^{1,21–23} and human subjects in both health and in the setting of skin disease^{2,4,5,24,25}. We demonstrate that after VACV viral infection through the skin, $CD8^+ T_{RM}$ cells are generated and distribute not only to the site of infection but also throughout the entire skin surface. These $CD8^+ T_{RM}$ cells produce effector cytokines, persist for many months, and are highly effective at rapidly controlling subsequent VACV skin infection. T_{RM} cells were orders of magnitude more effective than T_{CM} cells at controlling viral re-infection of the skin, at all time points examined in this study. The use of parabiotic mice allowed us to examine rigorously the tissue distribution and relative roles of T_{CM} and T_{RM} cells in VACV immune responses. The relatively minor role of $CD4^+$ T cells in VACV skin infection may reflect differences in immune responses to different viruses and/or infection of different tissues, as $CD4^+$ T cells are clearly more important in HSV infection^{12,26}. Moreover, $CD4^+$ T_{RM} cells predominate in human skin²⁴ and lung²⁵, and are enriched for memory cells specific for pathogens encountered through those tissues^{25,27,28}.

Pathogens typically invade the host through epithelial interfaces with the environment. Our studies suggest that T_{EM} cells generated as a result of epithelial tissue infections accumulate as T_{RM} cells at both sites of infection as well as at distant sites within the same epithelial tissue, providing broad and long-lived protective T-cell immunity against

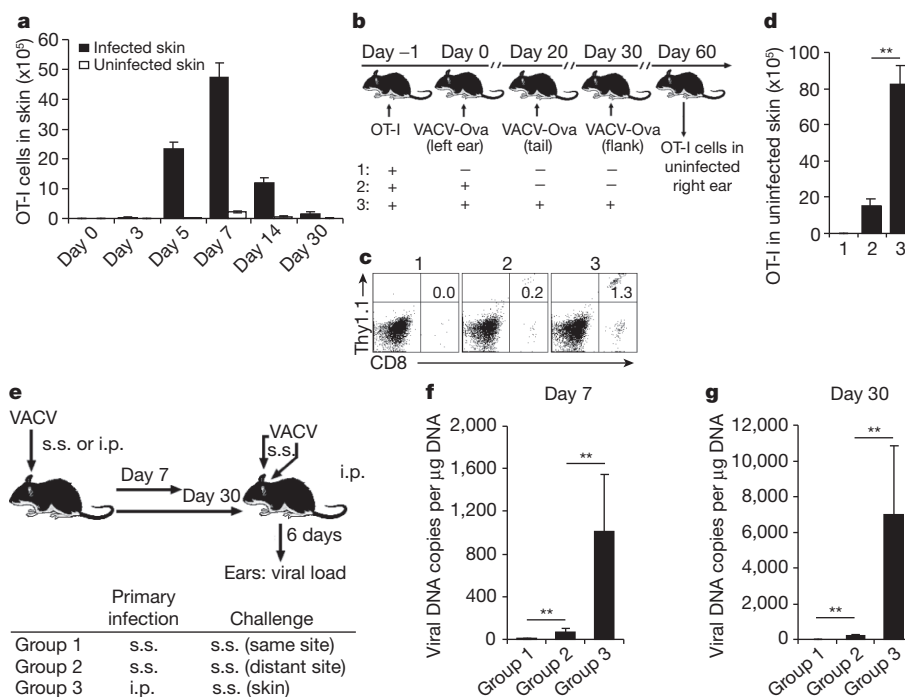


Figure 4 | Skin $CD8^+ T_{RM}$ cells also accumulate in unimmunized site after skin infection and are highly effective at eliminating virus. **a**, 2×10^6 Thy1.1⁺ OT-I cells were intravenously transferred into Thy1.2⁺ mice 1 day before 2×10^6 p.f.u. VACV-Ova skin scarification. The numbers of OT-I cells in infected and uninfected skin were enumerated over time. **b**, The left ears of OT-I-infused B6 mice were infected with VACV-Ova. Twenty and thirty days later, some of the mice were challenged with 2×10^6 p.f.u. VACV-Ova on the tail and flank skin, respectively. All infection routes were by skin scarification. Thirty days after challenge, Thy1.1⁺ OT-I cells in uninfected right ears were

examined. **c**, The percentages of Thy1.1⁺ OT-I cells in uninfected right ears. **d**, The numbers of Thy1.1⁺ OT-I cells in uninfected right ears. Error bars, s.e.m. **e**, μ MT mice were infected with VACV either by skin scarification on one ear or i.p. injection. Seven or thirty days later, mice were challenged with VACV on both ears via skin scarification (s.s.). Daily i.p. injection of FTY720 was performed. **f**, **g**, Six days later, viral loads in both infected and uninfected ears were measured. The mean and s.d. from triplicate qPCR runs are plotted. Data are representative of three independent experiments ($n = 5$ mice per group). ** $P < 0.01$.

re-infection. A more complete understanding of T_{RM}-cell-mediated immune memory should enhance our understanding of adaptive immunological memory, influence rational vaccine design, and illuminate the pathophysiology of human T-cell-mediated diseases.

METHODS SUMMARY

VACV skin infection. 2×10^6 p.f.u. of recombinant VACV Western Reserve strain or VACV expressing full-length ovalbumin (VACV-Ova) was used for epicutaneous infection by skin scarification or i.p. injection, as described previously²⁹.

Parabiotic mice. Well-matched (sex and age) mouse partners were anaesthetized to full muscle relaxation with ketamine and xylazine ($10 \mu\text{g g}^{-1}$) or 2.5% avertin ($15 \mu\text{l g}^{-1}$) by i.p. injection. The corresponding lateral aspects of mice were shaved and the excess hair was wiped off with alcohol prep pad. The disinfection was performed by wiping with betadine solution and 70% ethanol three times. Matching skin incisions were made from the olecranon to the knee joint of each mouse, and the subcutaneous fascia was bluntly dissected to create about 0.5 cm of free skin. The olecranon and knee joints were attached by a single 5-0 silk suture and tie, and the dorsal and ventral skins were approximated by staples or continuous suture. Betadine solution was used to cover the full length of the dorsal and ventral incision. The mice were then kept on heating pads and continuously monitored until recovery. For analgesic treatment, mice were injected subcutaneously with $2.5 \mu\text{g g}^{-1}$ flunixin every 8–12 h for 48 h after the operation.

Cell isolation from skin. Before harvest, skin hair was removed using Nair Hair Remover. Skin samples were then chopped into small fragments and incubated in Hanks balanced salt solution (HBSS) supplemented with 1 mg ml^{-1} collagenase A (Roche) and $40 \mu\text{g ml}^{-1}$ DNase I (Roche) at 37°C for 30 min. After filtrating through a $70\text{-}\mu\text{m}$ nylon cell strainer, cells were collected and washed thoroughly with cold PBS before staining.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 June 2011; accepted 12 January 2012.

Published online 29 February 2012.

- Liu, L. *et al.* Epidermal injury and infection during poxvirus immunization is crucial for the generation of highly protective T cell-mediated immunity. *Nature Med.* **16**, 224–227 (2010).
- Clark, R. A. *et al.* Skin effector memory T cells do not recirculate and provide immune protection in alemtuzumab-treated CTCL patients. *Sci. Transl. Med.* **4**, 117 (2012).
- Sheridan, B. S. & Lefrançois, L. Regional and mucosal memory T cells. *Nature Immunol.* **13**, 485–491 (2011).
- Boyman, O. *et al.* Spontaneous development of psoriasis in a new animal model shows an essential role for resident T cells and tumor necrosis factor- α . *J. Exp. Med.* **199**, 731–736 (2004).
- Conrad, C. *et al.* $\alpha_1\beta_1$ integrin is crucial for accumulation of epidermal T cells and the development of psoriasis. *Nature Med.* **13**, 836–842 (2007).
- Bevan, M. J. Helping the CD8⁺ T-cell response. *Nature Rev. Immunol.* **4**, 595–602 (2004).
- Antia, R., Ganusov, V. V. & Ahmed, R. The role of models in understanding CD8⁺ T-cell memory. *Nature Rev. Immunol.* **5**, 101–111 (2005).
- Kaech, S. M. & Wherry, E. J. Heterogeneity and cell-fate decisions in effector and memory CD8⁺ T cell differentiation during viral infection. *Immunity* **27**, 393–405 (2007).
- Lefrançois, L. & Obar, J. J. Once a killer, always a killer: from cytotoxic T cell to memory cell. *Immunol. Rev.* **235**, 206–218 (2010).
- Freyschmidt, E. J. *et al.* Skin inflammation arising from cutaneous regulatory T cell deficiency leads to impaired viral immune responses. *J. Immunol.* **185**, 1295–1302 (2010).
- Sanz, P. & Moss, B. Identification of a transcription factor, encoded by two vaccinia virus early genes, that regulates the intermediate stage of viral gene expression. *Proc. Natl Acad. Sci. USA* **96**, 2692–2697 (1999).
- Nakanishi, Y., Lu, B., Gerard, C. & Iwasaki, A. CD8⁺ T lymphocyte mobilization to virus-infected tissue requires CD4⁺ T-cell help. *Nature* **462**, 510–513 (2009).
- Wakim, L. M., Waithman, J., van Rooijen, N., Heath, W. R. & Carbone, F. R. Dendritic cell-induced memory T cell activation in nonlymphoid tissues. *Science* **319**, 198–202 (2008).
- Masopust, D., Vezys, V., Marzo, A. L. & Lefrançois, L. Preferential localization of effector memory cells in nonlymphoid tissue. *Science* **291**, 2413–2417 (2001).
- Lefrançois, L. Development, trafficking, and function of memory T-cell subsets. *Immunol. Rev.* **211**, 93–103 (2006).
- Klonowski, K. D. *et al.* Dynamics of blood-borne CD8 memory T cell migration *in vivo*. *Immunity* **20**, 551–562 (2004).
- Gebhardt, T. *et al.* Different patterns of peripheral migration by memory CD4⁺ and CD8⁺ T cells. *Nature* **477**, 216–219 (2011).
- Chong, B. F., Murphy, J. E., Kupper, T. S. & Fuhlbrigge, R. C. E-selectin, thymus- and activation-regulated chemokine/CCL17, and intercellular adhesion molecule-1 are constitutively coexpressed in dermal microvessels: a foundation for a cutaneous immunosurveillance system. *J. Immunol.* **172**, 1575–1581 (2004).
- Weninger, W. *et al.* Specialized contributions by $\alpha(1,3)$ -fucosyltransferase-IV and FucT-VII during leukocyte rolling in dermal microvessels. *Immunity* **12**, 665–676 (2000).
- Jiang, X., Campbell, J. J. & Kupper, T. S. Embryonic trafficking of $\gamma\delta$ T cells to skin is dependent on E/P selectin ligands and CCR4. *Proc. Natl Acad. Sci. USA* **107**, 7443–7448 (2010).
- Gebhardt, T. *et al.* Memory T cells in nonlymphoid tissue that provide enhanced local immunity during infection with herpes simplex virus. *Nature Immunol.* **10**, 524–530 (2009).
- Masopust, D. *et al.* Dynamic T cell migration program provides resident memory within intestinal epithelium. *J. Exp. Med.* **207**, 553–564 (2010).
- Wakim, L. M., Woodward-Davis, A. & Bevan, M. J. Memory T cells persisting within the brain after local infection show functional adaptations to their tissue of residence. *Proc. Natl Acad. Sci. USA* **107**, 17872–17879 (2010).
- Clark, R. A. *et al.* The vast majority of CLA⁺ T cells are resident in normal skin. *J. Immunol.* **176**, 4431–4439 (2006).
- Purwar, R. *et al.* Resident memory T cells (T_{RM}) are abundant in human lung: diversity, function, and antigen specificity. *PLoS ONE* **6**, e16245 (2011).
- Lund, J. M., Hsing, L., Pham, T. T. & Rudensky, A. Y. Coordination of early protective immunity to viral infection by regulatory T cells. *Science* **320**, 1220–1224 (2008).
- Román, E. *et al.* CD4 effector T cell subsets in the response to influenza: heterogeneity, migration, and function. *J. Exp. Med.* **196**, 957–968 (2002).
- Romani, L. Immunity to fungal infections. *Nature Rev. Immunol.* **11**, 275–288 (2011).
- Liu, L., Fuhlbrigge, R. C., Karibian, K., Tian, T. & Kupper, T. S. Dynamic programming of CD8⁺ T cell trafficking after live viral immunization. *Immunity* **25**, 511–520 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Tian, R. Purwar and Q. Zhan for technical assistance. We thank J. J. Campbell for discussion of the project. This work was supported by National Institutes of Health (NIH) grants R01AI041707, R37AI025082 and TR01 AI097128 to T.S.K.

Author Contributions X.J. and T.S.K. designed research; X.J. performed research; L.L. helped to establish the VACV skin scarification model; A.J.W. helped to create parabiotic mice; X.J., R.A.C., R.C.F., L.L. and T.S.K. analysed data; and X.J., R.A.C. and T.S.K. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.S.K. (tkupper@partners.org).

METHODS

Mice. C57BL/6, CD4^{-/-}, IFN- γ ^{-/-} and μ MT mice were purchased from The Jackson Laboratory. Thy1.1⁺ Rag^{-/-} OT-I, CD45.1⁺ OT-II, FucT IV/VII^{-/-} mice were housed at the animal facility of Harvard Institute of Medicine, Harvard Medical School. Thy1.1⁺ Rag^{-/-} OT-I mice were crossed with FucT IV/VII^{-/-} mice to yield Thy1.1⁺ FucT IV/VII^{-/-} OT-I mice. OT-I and OT-II are T-cell-receptor-transgenic mice recognizing chicken ovalbumin residues 257–264 in the context of H-2K^b and 323–339 in the context of I-A^b, respectively. Animal experiments were performed in accordance with the guidelines set out by the Center for Animal Resources and Comparative Medicine at Harvard Medical School.

Viruses and infections. Recombinant VACV (Western Reserve strain) and VACV-Ova were originally obtained from B. Moss (NIH) and grown on HeLa cells. Viral titres were determined using CV-1 cells and 2×10^6 p.f.u. of VACV or VACV-Ova was used for epicutaneous infection by skin scarification or i.p. injection, as described previously²⁹.

Parabiotic mice. Parabiosis surgery was performed as described elsewhere³⁰ with some modifications. Briefly, sex- and age-matched mouse partners were anaesthetized to full muscle relaxation with ketamine and xylazine ($10 \mu\text{g g}^{-1}$) or with 2.5% avertin ($15 \mu\text{g g}^{-1}$) by i.p. injection. The corresponding lateral aspects of mice were shaved and the excess hair was wiped off with alcohol prep pad. After skin disinfection by wiping with betadine solution and 70% ethanol three times, two matching skin incisions were made from the olecranon to the knee joint of each mouse, and the subcutaneous fascia was bluntly dissected to create about 0.5 cm of free skin. The olecranon and knee joints were attached by a single 5-0 silk suture and tie, and the dorsal and ventral skins were approximated by staples or continuous suture. Betadine solution was used to cover the full length of the dorsal and ventral incision. The mice were then kept on heating pads and continuously monitored until recovery. $2.5 \mu\text{g g}^{-1}$ flunixin was used for analgesic treatment by subcutaneous injection every 8–12 h for 48 h after the operation. After an interval of the indicated weeks, parabiotic mice were surgically separated by a reversal of the above procedure for the next experiments.

Adoptive transfer and T-cell depletion. Lymph nodes were collected from the naive female Thy1.1⁺ Rag^{-/-} OT-I, CD45.1⁺ OT-II, or Thy1.1⁺ FucT IV/VII^{-/-} OT-I mice at the age of 6–8 weeks. OT-I or OT-II cells were purified by magnetic cell sorting using mouse CD8 α ⁺ or CD4⁺ T-cell isolation kit (Miltenyi Biotec), respectively. 2×10^6 isolated OT-I and/or OT-II cells were then intravenously transferred to female recipient mice. In some experiments, OT-I and OT-II cells were labelled with carboxyfluorescein succinimidyl ester (CFSE) before co-transfer. To deplete CD4⁺ T cells *in vivo*, the recipient mice were injected i.p. with 500 μg anti-CD4 (GK1.5) in 100 μl PBS 4 days and 1 day before and on day 2 and 5 after infection.

Preparation of cell suspensions. Lymph nodes and spleen were harvested and mashed through a 70- μm nylon cell strainer to prepare cell suspensions. Red blood cells were lysed using lysing buffer. Skin tissue was removed after hair removal, chopped into small fragments and incubated in Hanks balanced salt solution (HBSS) supplemented with 1 mg ml^{-1} collagenase A and 40 $\mu\text{g ml}^{-1}$ DNase I at 37 °C for 30 min. After filtrating through a 70- μm nylon cell strainer, cells were collected and washed thoroughly with cold PBS before staining.

Intracellular cytokine detection. The infected skin from memory OT-I-bearing mice was harvested at 35 days after infection, and single-cell suspensions were prepared as described above. In some cases, 2 months after skin scarification with VACV-Ova, mice transferred with 2×10^6 OT-I cells were i.p. challenged with VACV-Ova. Five days after challenge, splenocytes were prepared. Red blood cells

were lysed using lysing buffer. Cells were then incubated with 2 $\mu\text{g ml}^{-1}$ SINFEKL peptide of ovalbumin at the presence of Brefeldin A for 7 h. Fc receptors were blocked with CD16/CD32 monoclonal antibodies and intracellular IFN- γ , TNF- α and IL-2 staining was performed using Intracellular Cytokine Detection Kits (BD Bioscience) before flow cytometry.

Determination of viral load. Mice were challenged with 2×10^6 p.f.u. VACV or VACV-Ova on the skin. In some cases, mice were simultaneously injected (i.p.) with 1 $\mu\text{g g}^{-1}$ FTY720 each day. At indicated time points, viral load of skin was examined by quantitative real-time PCR, as described previously²⁹.

Isolation of mRNA and real-time PCR. Total RNA was extracted from homogenized skin tissue and cDNA was generated with iScript cDNA synthesis kit (Bio-Rad). Bio-Rad iCycler iQ Real-Time PCR Detection System (Bio-Rad) was used with the following settings: 45 cycles of 15 s of denaturation at 95 °C, and 1 min of primer annealing and elongation at 60 °C. Real-time PCR was done with 1 μl cDNA plus 12.5 μl of 2 \times iQ SYBR Green Supermix (Bio-Rad) and 0.5 μl (10 μM) specific primers: mouse E-selectin 1 (5'-GGACACCACAAATCCCAGTCTG-3') and mouse E-selectin 2 (5'-TCGCAGGAGAACTCACAACCTGG-3'); mouse P-selectin 1 (5'-AAGATGCCTGGCTACTGGACAC-3') and mouse P-selectin 2 (5'-CAAGAGGCTGAACGCAGGTCAT-3'); mouse β -actin 1 (5'-CA TTGCTGACAGGATGCAGAAGG-3') and mouse β -actin 2 (5'-TGCTGGAA GGTGGACAGTGAGG-3'). All samples were run in duplicate and fold change of gene expression was calculated using the reference sample (naive skin).

Antibodies and flow cytometry. The following anti-mouse antibodies were obtained from BD PharMingen: CD8 α (53-6.7), Thy1.1 (OX-7), CD4 (L3T4), CD45.1 (A20), CD19 (1D3), CD16/CD32 (2.4G2), CD44 (IM7), CD62L (MEL-14), CD69 (H1.2F3), CD103 (M290), CD122 (TM-Beta 1), BCL2 (A19-3), $\alpha\beta$ 7 (DATK32), IFN- γ (XMG1.2), TNF- α (MP6-XT22), IL-2 (JES6-5H4). Fluorescence-conjugated anti-mouse CD127 (A7R34) and PD-1 (RMP1-30) were purchased from eBioscience. PE-conjugated B8R₂₀₋₂₇/H-2K^b pentamers were obtained from ProImmune Ltd and pentamer⁺ CD8⁺ T-cell staining was performed according to the protocol provided by the company. E- or P-selectin ligand expression was examined by incubating cells with rmE-Selectin/Fc Chimera or rmP-Selectin/Fc Chimera (R&D System) in conjunction with APC-conjugated F(ab')₂ fragments of goat anti-human IgG F(c) antibody (Jackson Immuno-research). Dead cells were excluded using 7-AAD staining. Data were analysed using FACSCanto Flow Cytometer using FACSDiva software.

Immunofluorescence microscopy. One centimetre of tail containing the infected skin site from OT-I-bearing mice was harvested and deboned at 45 days after skin scarification with VACV-Ova. The skin was collected and embedded in OCT (Tissue-Tek; Sakura) and frozen at -80 °C. Skin sections were performed on a cryostat (Leica CM1850 UV) at 6- μm thickness and air-dried for 6–8 h. Sections were then fixed in -20 °C acetone for 5 min, rehydrated with PBS, and blocked with 2% FCS in PBS for 15 min at room temperature (20 °C). After staining with Thy1.1-PE (1:200) in PBS for 1 h at room temperature, sections were rinsed three times (for 5 min each time) with TBS-tween 20 by shaking and mounted with DAPI Fluoromount-G (Southern Biotech). Images were acquired with a Nikon Eclipse E600 microscope and processed using SPOTSOFTWARE (version 4.6, Diagnostic Instruments).

Statistical analysis. Statistical significance in values between experimental groups was determined by one-way analysis of variance (ANOVA) followed by Tukey post-test. $P < 0.05$ was considered statistically significant.

30. Wagers, A. J., Sherwood, R. I., Christensen, J. L. & Weissman, I. L. Little evidence for developmental plasticity of adult hematopoietic stem cells. *Science* **297**, 2256–2259 (2002).